

**NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
THE SUCROSE PATHWAY**

**CROSS REFERENCE TO RELATED APPLICATIONS**

This application claims priority under 35 U.S.C §119(e) and/or 35 U.S.C §120 of applications No. 60/067,000 filed November 24, 1997; No. 60/069,472 filed December 9, 1997; No. 60/072,888 filed January 27, 1998; No. 60/074,201 filed February 10, 1998; No. 60/074,282 filed February 10, 1998; No. 60/074,280 filed February 10, 1998; No. 60/074,281 filed February 10, 1998; No. 60/074,566 filed February 12, 1998; No. 60/074,567 filed February 12, 1998; No. 60/074,565 filed February 12, 1998; No. 60/075,462 filed February 19, 1998; No. 60/074,789 filed February 19, 1998; No. 60/075,459 filed February 19, 1998; No. 60/075,461 filed February 19, 1998; No. 60/075,464 filed February 19, 1998; No. 60/075,460 filed February 19, 1998; No. 60/075,463 filed February 19, 1998; No. 60/076,912 filed March 6, 1998; No. 60/077,231 filed March 9, 1998; No. 60/077,229 filed March 9, 1998; No. 60/077,230 filed March 9, 1998; No. 60/078,368 filed March 18, 1998; No. 60/080,844 filed April 7, 1998; No. 60/083,067 filed April 27, 1998, "Nucleic Acid Molecules and Other Molecules Associated with Plants.(soymon016)" docket No. 38-21(15348)A filed April 29, 1998; No. 60/083,387 filed April 29, 1998; No. 60/083,388 filed April 29, 1998; No. 60/083,389 filed April 29, 1998, "Nucleic Acid Molecules and Other Molecules Associated with the Phosphogluconate Pathway." docket No. 38-21(15365)A filed April 30, 1998; No. 60/085,224 filed May 13, 1998; No. 60/085,223 filed May 13, 1998; No. 60/085,222 filed May 13, 1998; No. 60/086,186 filed May 21, 1998; No. 60/086,187 filed May 21, 1998; No. 60/086,185 filed May 21, 1998; No. 60/086,184 filed May 21, 1998; No. 60/086,183 filed May 21, 1998; No. 60/086,188 filed May 21, 1998; No. 60/087,422 filed June 1, 1998; No. 60/089,524 filed June 16, 1998; No. 60/089,810 filed June

669210"EST.250

18, 1998; No. 60/089,814 filed June 18, 1998; No. 60/089,793 filed June 18, 1998; No. 60/090,170 filed June 22, 1998; No. 60/090,928 filed June 26, 1998; No. 60/091,035 filed June 29, 1998; No. 60/091,405 filed June 30, 1998; No. 60/092,036 filed July 8, 1998; No. 60/099,667 filed September 9, 1998; No. 60/099,670 filed September 9, 1998; No. 60/099,697 filed September 9, 1998; No. 60/100,674 filed September 16, 1998; No. 60/100,673 filed September 16, 1998; No. 60/100,672 filed September 16, 1998; No. 60/101,131 filed September 21, 1998; No. 60/101,132 filed September 21, 1998; No. 60/101,130 filed September 21, 1998; No. 60/101,508 filed September 22, 1998; No. 60/101,344 filed September 22, 1998; No. 60/101,347 filed September 22, 1998; No. 60/101,343 filed September 22, 1998; No. 60/101,707 filed September 25, 1998; No. 60/104,126 filed October 13, 1998; No. 60/104, 128 filed October 13, 1998; No. 60/104,127 filed October 13, 1998; No. 60/104,124 filed October 13, 1998; No. 60/104,123 filed October 13, 1998; No. 60/109,018 filed November 18, 1998; No. 60/108,996 filed November 18, 1998, "Nucleic Acid Molecules and Other Molecules Associated With Plants" docket No. 38-21(15075)B filed November 24, 1998; No. 09/210,297 filed December 8, 1998, "Nucleic acid Molecules and other Molecules associated with Plants" docket No. 38-21(15668)A filed December 11, 1998; No. 60/113,224 filed December 22, 1998 and "Nucleic Acid Molecules and Other Molecules Associated with Transcription in Plants" docket No. 38-21(15300)B filed January 12, 1999, all of which are herein incorporated by reference in their entirety.

### **FIELD OF THE INVENTION**

The present invention is in the field of plant biochemistry. More specifically the invention relates to nucleic acid sequences from plant cells, in particular, nucleic acid sequences from maize and soybean plants associated with the sucrose pathway. The invention encompasses

nucleic acid molecules that encode proteins and fragments of proteins. In addition, the invention also encompasses proteins and fragments of proteins so encoded and antibodies capable of binding these proteins or fragments. The invention also relates to methods of using the nucleic acid molecules, proteins and fragments of proteins and antibodies, for example for genome mapping, gene identification and analysis, plant breeding, preparation of constructs for use in plant gene expression and transgenic plants.

### **BACKGROUND OF THE INVENTION**

Carbon fixed during photosynthesis is either retained in the chloroplast and converted to a storage carbohydrate, for example, starch, or it is transferred to the cytosol in the form of triose phosphates and converted to sucrose. The newly synthesized sucrose in source tissues is a major transported form of reduced carbon in higher plants and can be either metabolized into other carbohydrates, stored in the vacuole or exported to other plant tissues. Plant tissues where sucrose is synthesized, such as leaves, are often referred to as 'source' tissues. Translocated sucrose is retained in 'sink' tissues (such as expanding leaves, growing seeds, flowers, roots or tubers, and fruit) and may be assimilated, or further metabolized to sustain cell maintenance or fuel growth, or be converted to alternative storage compounds (*e.g.*, starch, fats). The relative type and size of these carbohydrate pools vary during tissue development, between different plant species, and within the same species subject to different environmental conditions. Such differences are reported to affect the yield and quality of agricultural produce.

Sucrose synthesis and catabolism are reported to be highly coordinated and regulated processes that may also be coordinately regulated with other dedicated metabolic pathways in a particular plant, plant organ or cell type. Sucrose synthesis is reported to be coordinately regulated with starch metabolism and photosynthesis in green 'source' plant tissues. Sucrose

supply by transport mechanisms to actively growing 'sink' tissues is reported to be coordinated with plant development. In growing sink tissues, the supply of carbohydrate is reported to be important to other metabolic pathways and physiological processes including respiration, starch biosynthesis, cell wall biogenesis, lipid and protein biosynthesis. Sucrose synthesis and/or transport is also reported to play a role in the carbohydrate capacity that is available to growing fruits and seeds. Sucrose resynthesis during seed germination is reported to play a role in seedling vigor and agronomic stand establishment in many plant species during early plant development.

In many plant species, enzymes of pathways involved in sucrose metabolism can play a role in plant physiology and plant growth and development. Compartmentation and temporal regulation of genes and enzymes of sucrose metabolic pathways can allow multiple pathways to utilize sucrose as a common metabolite. Flux through a particular sucrose metabolic pathway can define the utilization of sucrose in any tissue or developmental stage. Sucrose and its metabolite products have been reported to play a role in gene regulation and expression of the sucrose pathway and other metabolic pathways in plants.

Reviews on sucrose metabolism in plants include Avigad, In: *Encyclopedia of Plant Physiology*, Vol 13A, Loewus and Tanner, eds., Springer Verlag, Heidelberg, 217-347 (1982); Hawker, In: *Biochemistry of Storage Carbohydrates in Green Plants*, Dey and Dixon, eds., Academic Press, London, 1-51 (1985); Huber *et al.*, In: *Carbon Partitioning Within and Between Organisms*, Pollock *et al.*, eds., Bios Scientific, Oxford, 1-26 (1992); Stitt *et al.*, In: *Biochemistry of Plants*, Vol 10, Hatch and Boardman, eds., Academic Press, New York, 327-407 (1987); Quick and Schaffer, In: *Photoassimilate Distribution In: Plants And Crops*, Zamski and



Schaffer, eds., Marcel Dekker Inc., New York, 115-156 (1996), all of which are herein incorporated by reference in their entirety.

The synthesis of sucrose precursors (triose and hexose phosphates) is derived from either photosynthetic CO<sub>2</sub> fixation or degradation of previously deposited storage reserves. One substrate for sucrose synthesis in photosynthetic tissues is three carbon sugar phosphates. These are exported from the chloroplast during photosynthesis, predominantly in the form of triose phosphates. The pool of triose phosphates, dihydroxyacetone phosphate ("DHAP"), and glyceraldehyde-3-phosphate ("GAP"), is maintained at equilibrium within the cytoplasm by triose phosphate isomerase (EC 5.3.1.1). A subsequent reaction involves an aldol condensation of DHAP and GAP, catalyzed by the enzyme fructose 1,6-bisphosphate aldolase (often called aldolase) (EC 4.1.2.13) to form fructose 1,6-bisphosphate ("F1,6BP"). Fructose-1,6-bisphosphatase ("FBPase") (EC 3.1.3.11) catalyzes the cleavage of phosphate from the C1 carbon of fructose-1,6-bisphosphate to form fructose-6-phosphate ("F6P"). This reaction is essentially irreversible and has been reported to represent the first committed step within the pathway of sucrose synthesis. The cytosolic FBPase has been reported to be subject to allosteric regulation and may serve to coordinate the rate of sucrose synthesis with that of photosynthesis. Fructose 2,6-bisphosphate ("F2,6BP") is reported to be a regulator of FBPase (Black *et al.*, In: *Regulation of Carbohydrate Partitioning In Photosynthetic Tissue*, Heath and Preiss, eds., Waverly, Baltimore, 109-126 (1985); Stitt *et al.*, In: *Biochemistry Of Plants*, Vol. 10, Hatch and Boardman, eds., Academic Press, New York, 327-407 (1987), both of which are herein incorporated by reference in their entirety). The concentration of F2,6BP is reported to be controlled in plants by two enzymes, fructose-2,6-bisphosphatase (F2,6BPase) (EC 3.1.3.46) and fructose-6-phosphate,2-kinase (F6P,2K) (EC 2.7.1.105) (Stitt, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 41: 153-181 (1990), the entirety of which is herein incorporated by reference).

Glucose-6-phosphate ("G6P") and glucose-1-phosphate ("G1P") are reported to be maintained in equilibrium with the F6P pool by the action of phosphoglucoisomerase ("PGI")

(EC 5.3.1.9) and phosphoglucomutase ("PGM") (EC 5.4.2.2), respectively. Uridine diphosphate glucose ("UDPG") and pyrophosphate ("PPi") are formed from uridine triphosphate ("UTP") and G1P catalyzed by the enzyme UDPG-pyrophosphorylase ("UDPGase") (EC 2.7.7.9). This reaction is reversible and net flux in the direction of sucrose synthesis is reported to require removal of its products, particularly PPi. A pyrophosphate-dependent proton pump, vacuolar H<sup>+</sup>-translocating-pyrophosphatase (EC 3.6.1.1), has been identified within the vacuolar membrane and has been reported to utilize pyrophosphate to sustain a proton gradient formed between these two compartments (Rea *et al.*, *Trends in Biol. Sci.* 17: 348-353 (1992), the entirety of which is herein incorporated by reference).

A pyrophosphate-dependent fructose-6-phosphate phosphotransferase ("PFP") (EC 2.7.1.90) is also present in the cytoplasm and catalyzes the reversible production of F1,6BP and Pi from F6P and PPi. One reported function of PFP is to operate in a futile cycle with the cytosolic FBPase, and function as a "pseudopyrophosphatase" recycling PPi. Uridine diphosphate glucose is then combined with F6P to form sucrose-6-phosphate ("S6P"). This reaction is catalyzed by sucrose phosphate synthase ("SPS") (EC 2.4.1.14). Attachment of UDP to the glucose moiety activates the C1 carbon atom of UDPG, which is necessary for the subsequent formation of a glycosidic bond in sucrose. In certain organisms, SPS is capable of using adenine diphosphate glucose ("ADPG"), instead of UDPG, as a substrate. The use of nucleotide biphosphate sugars is a feature of metabolic pathways leading to the production of disaccharides and polysaccharides. SPS is reported to be subject to allosteric and covalent regulation and, in conjunction with the cytosolic FBPase, reportedly serves to coordinate the rate of sucrose synthesis with the rate of photosynthesis. The reported final reaction in the pathway is catalyzed by sucrose-6-phosphate phosphatase ("SPPase" or "SPP") (EC 3.1.3.24), which catalyzes the hydrolysis of S6P to sucrose. It has been reported that SPS and SPPase may associate to form a multienzyme complex, that the rate of sucrose-6-phosphate synthesis by SPS is enhanced in the presence of SPP, and that the rate of sucrose-6-phosphate hydrolysis by SPP is

increased in the presence of SPS (Echeverria *et al.*, *Plant Physiol.* 115: 223-227 (1997), herein incorporated by reference in its entirety).

## I. SUCROSE SYNTHESIS

Reviews describing fructose-1,6-bisphosphatase ("FBPase", EC 3.1.3.11) include those by Hers and Van Schaftingen, *Biochem J.* 206:1-12 (1982), the entirety of which is herein incorporated by reference, and Stitt, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 41:153-181 (1990). Two isoforms of FBPase are reported to exist in plants. The first isoform is associated with the plastid and occurs largely in photosynthetic plastids. The second isoform, located in the cytoplasm, is reported to be involved in both gluconeogenesis and sucrose synthesis (Zimmerman *et al.*, *J. Biol. Chem.* 253: 5952-5956 (1978); Stitt and Heldt, *Planta* 164: 179-188 (1985), both of which are hereby incorporated by reference in their entirety). FBPase catalyzes an irreversible reaction in the direction of F6P synthesis *in vivo* and has been reported to represent the first committed step in the pathway of sucrose synthesis. The properties of the enzyme are reported to involve the action of several regulatory metabolites (Stitt *et al.*, In: *Biochemistry Of Plants*, Vol. 10, Hatch and Boardman, eds., Academic Press, New York, 327-407 (1987)). The enzyme reportedly has a high affinity for its substrate F1,6BP, a requirement for  $Mg^{2+}$ , a requirement for a neutral pH, is weakly inhibited ( $K_m$  2-4  $\mu M$ ) by adenosine monophosphate (AMP), and is strongly inhibited by the regulatory metabolite F2,6BP (Hers and Van Schaftingen, *Biochem J.* 206: 1-12 (1982); Black *et al.*, In: *Regulation of Carbohydrate Partitioning In Photosynthetic Tissue*, Heath and Preiss, eds., Waverly, Baltimore, 109-126 (1985); Huber, *Annu. Rev. Plant Physiol.* 37: 233-246 (1986); Stitt *et al.*, In: *Biochemistry Of Plants*, Vol. 10, Hatch and Boardman, eds., Academic Press, New York, 327-407 (1987), all of which are herein incorporated by reference in their entirety). F2,6BP is also an activator of PFP and reportedly plays a role in the regulation of gluconeogenic and respiratory metabolism.

The concentration of F2,6BP is reportedly determined in plants by two enzymes, fructose-2,6-bisphosphatase ("F2,6BPase") (EC 3.1.3.46) and fructose-6-phosphate,2-kinase

("F6P,2K") (EC 2.7.1.105). A review of these enzymes is provided by Stitt, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 41: 153-181 (1990). Regulation of the activity of the F1,6FBPase and the rate of sucrose synthesis is reported to be, at least in part, brought about by changes in the concentration of F2,6BP.

Sucrose phosphate synthase (SPS (EC 2.4.1.14)) catalyzes a reaction that is displaced from equilibrium *in vivo* in the direction of S6P synthesis and is reported as an essentially irreversible reaction *in vivo* (Stitt *et al.*, In: *Biochemistry Of Plants*, Vol. 10, Hatch and Boardman, eds., Academic Press, New York, 327-407 (1987); Lunn and Rees, *Biochem. J.* 267: 739-743 (1990), the entirety of which is herein incorporated by reference; U.S. Patent No. 5,665,892, the entirety of which is herein incorporated by reference). SPS has been purified from spinach and *Zea mays*, and the amino acid and cDNA sequences have been published (Worrel *et al.*, *Plant Cell* 3:1121-1130 (1991); Klein *et al.*, *Planta* 190: 498-510 (1993); Sonnewald *et al.*, *Planta* 189: 174-181 (1993), all of which are herein incorporated by reference in their entirety). The enzyme has a subunit molecular weight of 117 kDa from spinach (Klein *et al.*, *Planta* 190: 498-510 (1993); Sonnewald *et al.*, *Planta* 189: 174-181 (1993), both of which are herein incorporated by reference) and pea (Lunn and Rees, *Phytochem.* 29: 1057-1063 (1990), the entirety of which is herein incorporated by reference) and 135 kDa from *Zea mays* (Worrel *et al.*, *Plant Cell* 3:1121-1130 (1991)). The native enzyme reportedly exists as a tetramer (Walker and Huber, *Plant Physiol.* 89: 518-524 (1988); Lunn and Rees, *Phytochem.* 29: 1057-1063 (1990); Worrel *et al.*, *Plant Cell* 3:1121-1130 (1991), although dimeric molecular weights have been reported (Klein *et al.*, *Planta* 190: 498-510 (1993), the entirety of which is herein incorporated by reference). Activity has been observed for SPS at both dimeric and tetrameric molecular weights (Sonnewald *et al.*, *Planta* 189:174-181 (1993), the entirety of which is herein incorporated by reference).

SPS is located in the cytosol, has a neutral pH optimum, and has been detected in all plant tissues which undertake active sucrose synthesis. SPS is also reported to undertake active sucrose synthesis. An increase in abundance of the enzyme is has been reported during the

development of leaves, germination of seeds and ripening of fruit. The enzyme has been reported to be subject to regulation by metabolites and is activated by G6P and is inhibited by Pi. Pi and GP6 are reported to act competitively at an allosteric site of the enzyme. In the presence of high Pi concentrations, the enzyme is phosphorylated which reduces activity of the enzyme. It has also been reported that light-induced photosynthesis increases the activity of SPS in crude extracts (Sicher and Kremer, *Plant Physiol.* 79: 910-912 (1984), Sicher and Kremer, *Plant Physiol.* 79: 695-698 (1985); Pollock and Housley, *Ann. Bot.* 55: 593-596 (1985), all of which are herein incorporated by reference in their entirety). In addition, it has been reported that compounds altering the phosphate status of the leaf can simulate the effects of light. Feeding leaves mannose, which sequesters phosphate by its conversion to the non-metabolized mannose-6-P, has been reported to cause activation of SPS (Stitt *et al.*, *Planta* 174: 217-230 (1988), the entirety of which is herein incorporated by reference).

The phosphorylation and dephosphorylation of SPS is catalyzed by SPS-phosphatase and SPS-kinase, respectively (Huber *et al.*, *Plant Physiol.* 99: 1275-1278 (1992). Hydrolysis of sucrose-6-P to sucrose is catalyzed by sucrose-6-phosphatase (SPPase or SPP) (EC 3.1.3.24). The activity of both SPS and SPP is reported to be affected by a multienzyme complex between SPS and SPP (Echeverria *et al.*, *Plant Physiol.* 115: 223-227 (1997)).

Regulatory properties of SPS and FBPase are reported to coordinate the rate of sucrose synthesis with that of photosynthesis (Stitt, In: *Plant Physiology, Biochemistry and Molecular Biology*, Dennis and Turpin, eds., Singapore, London, 319-340 (1990), the entirety of which is herein incorporated by reference). When photosynthesis produces triose phosphate in excess of the rate of sucrose synthesis, a feed-forward activation of sucrose synthesis occurs. Triose phosphate crosses the chloroplast membrane in exchange for cytosolic Pi. Under these conditions, F6P,2-kinase activity is reduced and the inhibition of F2,6BPase is decreased.

As cytosolic F2,6BP falls, F2,6BPase activity increases, and F6P levels increase. Hexose phosphate levels are reported to increase due to PGM and PGI, and with low Pi, activate SPS and F1,6BPase. Reduction in rate of photosynthesis must result in a deactivation of sucrose

synthesis, which occurs through decreased cytosolic triose-P, increased Pi and ultimately increased F2,6BP concentration and reduced SPS activity (Stitt, *Phil. Trans. R Soc. Lond. B* 342: 225-233 (1993); Huber *et al.*, *Plant Physiol.* 99: 1275-1278 (1992); Neuhaus *et al.*, *Planta* 181: 583-592 (1990), both of which are herein incorporated by reference).

## II. METABOLIC PATHWAYS OF SUCROSE CATABOLISM

Sucrose can initially be cleaved by invertases (EC 3.2.1.26) or by sucrose synthases (EC 2.4.1.13). Invertases, which are classified as acid or alkaline in pH preference (Karuppiiah *et al.*, *Plant Physiol.* 91: 993-998 (1989); Fahrendorf and Beck, *Planta* 180: 237-244 (1990); Iwatsubo *et al.*, *Biosc. Biotech. Biochem.* 56: 1959-1962 (1992); Unger *et al.*, *Plant Physiol.* 104: 1351-1357 (1994); Avigad, In: *Encyclopedia of Plant Physiology*, Vol 13A, Loewus and Tanner, eds., Springer Verlag, Heidelberg, 217-347 (1982), all of which are herein incorporated by reference in their entirety), irreversibly cleave sucrose into glucose and fructose, both of which is usually phosphorylated for further metabolism. The invertase pathway usually is associated with rapidly growing sink tissues such as expanding leaves, expanding internodes, flower petals, and early fruit development (Avigad, In: *Encyclopedia of Plant Physiology*, Vol 13A, Loewus and Tanner, eds., Springer Verlag, Heidelberg, 217-347 (1982); Huber, *Plant Physiol.* 91: 656-662 (1989); Morris and Arthur, *Phytochem.* 23: 2163-2167 (1984); Hawker *et al.*, *Phytochem.* 15: 1441-1443 (1976); Schaffer *et al.*, *Plant Physiol.* 69: 151-155 (1987), all of which are herein incorporated by reference in their entirety).

Sucrose synthase carries out the kinetically reversible transglycosylation of sucrose and UDP into fructose and UDPG, requiring only the phosphorylation of fructose for additional metabolism. Polysaccharide biosynthesis in sink tissues may utilize a sucrose synthase mediated sucrose catabolism (Avigad, In: *Encyclopedia of Plant Physiology*, Vol 13A, Loewus and Tanner, eds., Springer Verlag, Heidelberg, 217-347 (1982); Doehlert *et al.*, *Plant Physiol.* 86: 1013-1019 (1988); Dale and Housley *Plant Physiol.* 82: 7-10 (1986), all of which are herein

incorporated by reference). Respiring tissues reportedly utilize either sucrose synthase or invertase metabolic pathways (Echeverria and Humphreys, *Phytochem.* 23: 2173-2178 (1984); Uritani and Asahi, In: *The Biochemistry of Plants* Vol. 2, Davies, ed., Academic Press, New York, 463-487 (1980), all of which are herein incorporated by reference in their entirety). Tissues that are undergoing respiration, starch biosynthesis, amino acid and fatty acid synthesis, rapid expansion or growth, and other cellular metabolism, can utilize several sucrose metabolic pathways which may be temporally or compartmentally regulated (Doehlert *et al.*, *Plant Physiol.* 86: 1013-1019 (1988); Doehlert, *Plant Physiol.* 78: 560-567 (1990); Doehlert and Choury, In: *Recent Advances in Phloem Transport and Assimilate Compartmentation*, Bonnemain *et al.*, eds., Oest editions, Nantes, France, 187-195 (1991); Delmer and Stone, In: *The Biochemistry of Plants*, Vol. 14, Preiss, ed., Academic Press, San Diego, 373-420 (1988); Maas *et al.*, *EMBO J.* 9: 3447-3452 (1990), all of which are herein incorporated by reference in their entirety).

Hexose kinases are a class of enzymes responsible for the phosphorylation of hexoses, and are classified into two groups. Hexokinase (EC 2.7.1.1) can phosphorylate either glucose or fructose, with different isoforms often unique to different tissues or plant species. Different isoforms can have affinities for different hexoses (Turner and Copeland, *Plant Physiol.* 68: 1123-1127 (1981), the entirety of which is herein incorporated by reference; Copeland and Turner, In: *The Biochemistry of Plants*, Vol. 11, Stumpf and Conn, eds., Academic Press, New York, 107-128 (1987), the entirety of which is herein incorporated by reference). Hexokinases include fructokinases (EC 2.7.1.11), which typically have specific affinities for fructose (Doehlert, *Plant Physiol.* 89: 1042-1048 (1989); Renz and Stitt *Planta* 190: 166-175 (1993), both of which are herein incorporated by reference). Fructokinases can also be specific in their affinity for nucleotides. The extent to which a fructokinase utilizes UTP may play a physiological role in

how efficiently UDP can be recycled for sucrose synthase activity in a particular tissue (Huber and Akazawa, *Plant Physiol.* 81: 1008-1013 (1986); Xu *et al.*, *Plant Physiol.* 90: 635-642 (1989), both of which are herein incorporated by reference). UDP levels for the sucrose synthase reaction may be maintained, even in the case of an ATP-specific fructokinase, by the enzyme NDP-kinase (EC 2.7.4.6).

NDP-kinase has been reported in several plant tissues (Kirkland and Turner, *J. Biochem.* 72: 716-720 (1959); Bryce and Nelson, *Plant Physiol.* 63: 312-317 (1979); Dancer *et al.*, *Plant Physiol.* 92: 637-641 (1990); Yano *et al.*, *Plant Molec. Biol.* 23: 1087-1090 (1993), all of which are herein incorporated by reference in their entirety). Fructokinase can be substrate inhibited by fructose. In addition, sucrose synthase can be inhibited by fructose (Doehlert, *Plant Sci.* 52: 153-157 (1987); Morell and Copeland, *Plant Physiol.* 78: 140-154 (1985), Ross and Davies, *Plant Physiol.* 100: 1008-1013 (1992), all of which are herein incorporated by reference in their entirety). Whereas plant tissues where sucrose is catabolized by sucrose synthase predominantly contain fructokinases (Xu *et al.*, *Plant Physiol.* 90: 635-642 (1989); Kursanov *et al.*, *Soviet Plant Physiol.* 37: 507-515 (1990); Ross *et al.*, *Plant Physiol.* 90: 748-756 (1994)), plant tissues where sucrose is catabolized by invertase often contain hexokinases (Nakamura *et al.*, *Plant Physiol.* 81: 215-220 (1991)). Tissues which have both invertase and sucrose synthase activity may contain both hexose kinases (Nakamura *et al.*, *Plant Physiol.* 81: 215-220 (1991), the entirety of which is herein incorporated by reference). F6P resulting from hexose kinase activity can be further metabolized in glycolysis or used in resynthesis of sucrose by SPS. G6P resulting from hexose kinase activity can enter the pentose phosphate pathway, via G6P dehydrogenase (EC 1.1.1.49), or be converted to F6P by phosphoglucosomerase ("PGI") (EC 5.3.1.9) or G1P by phosphoglucomutase ("PGM") (EC 5.4.2.2) (Rees, In: *Encyclopedia of Plant Physiology* Vol 18,



Douce and Day, eds., Springer Verlag, Berlin, 391-417 (1985); Copeland and Turner, In: *The Biochemistry of Plants* Vol. 11, Stumpf and Conn, eds., Academic Press, New York, 107-128 (1987); Foster and Smith, *Planta* 180: 237-244 (1993), all of which are herein incorporated by reference in their entirety).

PGI and PGM are reported to be ubiquitous and reversible with commitments of G6P to either F6P or G1P resulting from fluxes in metabolites further along each pathway, *i.e.*, depending on the cell needs for glycolysis (F6P) or starch biosynthesis (G1P) (Edwards and Rees, *Phytochem.* 25: 2033-2039 (1986); Kursanov *et al.*, *Soviet Plant Physiol.* 37: 507-515 (1990); Tobias *et al.*, *Plant Physiol.* 99: 140-145 (1992), all of which are herein incorporated by reference in their entirety). UDPG formed by sucrose synthase may be utilized directly for cellulose or callose biosynthesis via UDP-glucose dehydrogenase (EC 1.1.1.2) (Robertson *et al.*, *Phytochem.* 39: 21-28 (1995), the entirety of which is herein incorporated by reference), can be used for sucrose synthesis by SPS or sucrose synthase, or for glycolysis or starch metabolism dependent on further metabolism by UDP-glucose pyrophosphorylase (EC 2.7.7.9). UDP-glucose phosphorylase has been reported to be a largely reversible enzyme (Kleczkowski, *Phytochem.* 37: 1507-1515 (1994), the entirety of which is herein incorporated by reference). Flux through UDP-glucose pyrophosphorylase is reported to be influenced by metabolite levels and utilization of reaction products further along in the pathways (Doehlert *et al.*, *Plant Physiol.* 86: 1013-1019 (1988); Huber and Akazawa, *Plant Physiol.* 81: 1008-1013 (1986); Zrenner *et al.*, *Planta* 190: 247-252 (1993), all of which are herein incorporated by reference in their entirety). The reversibility of PGI, PGM and UDPGPPase has been reported to provide for metabolic variability and networking in metabolism, independent of which initial enzyme cleaved sucrose.

69274306922260

The fate of F6P reportedly plays a role in carbohydrate metabolism. NTP-phosphofructokinase (PFK) (EC 2.7.1.11) (Copeland and Turner, In: *The Biochemistry of Plants* Vol. 11, Stumpf and Conn, eds., Academic Press, New York, 107-128 (1987); Dennis and Greyson, *Plant Physiol.* 69: 395-404 (1987); Rees, In: *The Biochemistry of Plants* Vol. 14, Preiss, ed., Academic Press, San Diego, 1-33 (1988), all of which are herein incorporated by reference in their entirety) is reported to irreversibly convert F6P to F16BP and is associated with glycolysis. The reverse reaction of F16BP to F6P, associated with gluconeogenesis, is essentially irreversible, and is catalyzed by FBPase (EC 3.1.3.11) (Black *et al.*, *Plant Physiol.* 69: 387-394 (1987). Both reactions may be carried out in a reversible manner by a PPi-dependent fructose-6-phosphate phosphotransferase or PPi-phosphofructokinase (PFP; EC 2.7.1.90) (Black *et al.*, *Plant Physiol.* 69: 387-394 (1987).

PPi-dependent fructose-6-phosphate phosphotransferase or PPi-phosphofructokinase is reported to play a role in the generation of biosynthetic intermediates (Dennis and Greyson, *Plant Physiol.* 69: 395-404 (1987); Tobias *et al.*, *Plant Physiol.* 99: 146-152 (1992), the entirety of which is herein incorporated by reference) in addition to the cycling of PPi for UDPGPPase and ultimately UDP for sucrose synthase (Huber and Akazawa, *Plant Physiol.* 81: 1008-1013 (1986); Black *et al.*, *Plant Physiol.* 69: 387-394 (1987); Rees, In: *The Biochemistry of Plants* Vol. 14, Preiss, ed., Academic Press, San Diego, 1-33 (1988), all of which are herein incorporated by reference in their entirety).

## II. EXPRESSED SEQUENCE TAG NUCLEIC ACID MOLECULES

Expressed sequence tags, or ESTs are randomly sequenced members of a cDNA library (or complementary DNA)(McCombie *et al.*, *Nature Genetics* 1:124-130 (1992); Kurata *et al.*, *Nature Genetics* 8:365-372 (1994); Okubo *et al.*, *Nature Genetics* 2:173-179 (1992), all of which

references are incorporated herein in their entirety). The randomly selected clones comprise insets that can represent a copy of up to the full length of a mRNA transcript.

Using conventional methodologies, cDNA libraries can be constructed from the mRNA (messenger RNA) of a given tissue or organism using poly dT primers and reverse transcriptase (Efstratiadis *et al.*, *Cell* 7:279-3680 (1976), the entirety of which is herein incorporated by reference; Higuchi *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 73:3146-3150 (1976), the entirety of which is herein incorporated by reference; Maniatis *et al.*, *Cell* 8:163-182 (1976) the entirety of which is herein incorporated by reference; Land *et al.*, *Nucleic Acids Res.* 9:2251-2266 (1981), the entirety of which is herein incorporated by reference; Okayama *et al.*, *Mol. Cell. Biol.* 2:161-170 (1982), the entirety of which is herein incorporated by reference; Gubler *et al.*, *Gene* 25:263-269 (1983), the entirety of which is herein incorporated by reference).

Several methods may be employed to obtain full-length cDNA constructs. For example, terminal transferase can be used to add homopolymeric tails of dC residues to the free 3' hydroxyl groups (Land *et al.*, *Nucleic Acids Res.* 9:2251-2266 (1981), the entirety of which is herein incorporated by reference). This tail can then be hybridized by a poly dG oligo which can act as a primer for the synthesis of full length second strand cDNA. Okayama and Berg, *Mol. Cell. Biol.* 2:161-170 (1982), the entirety of which is herein incorporated by reference, report a method for obtaining full length cDNA constructs. This method has been simplified by using synthetic primer-adapters that have both homopolymeric tails for priming the synthesis of the first and second strands and restriction sites for cloning into plasmids (Coleclough *et al.*, *Gene* 34:305-314 (1985), the entirety of which is herein incorporated by reference) and bacteriophage vectors (Krawinkel *et al.*, *Nucleic Acids Res.* 14:1913 (1986), the entirety of which is herein

incorporated by reference; Han *et al.*, *Nucleic Acids Res.* 15:6304 (1987), the entirety of which is herein incorporated by reference).

These strategies have been coupled with additional strategies for isolating rare mRNA populations. For example, a typical mammalian cell contains between 10,000 and 30,000 different mRNA sequences (Davidson, *Gene Activity in Early Development*, 2nd ed., Academic Press, New York (1976), the entirety of which is herein incorporated by reference). The number of clones required to achieve a given probability that a low-abundance mRNA will be present in a cDNA library is  $N = (\ln(1-P))/(\ln(1-1/n))$  where N is the number of clones required, P is the probability desired and 1/n is the fractional proportion of the total mRNA that is represented by a single rare mRNA (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press (1989), the entirety of which is herein incorporated by reference).

A method to enrich preparations of mRNA for sequences of interest is to fractionate by size. One such method is to fractionate by electrophoresis through an agarose gel (Pennica *et al.*, *Nature* 301:214-221 (1983), the entirety of which is herein incorporated by reference). Another such method employs sucrose gradient centrifugation in the presence of an agent, such as methylmercuric hydroxide, that denatures secondary structure in RNA (Schweinfest *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 79:4997-5000 (1982), the entirety of which is herein incorporated by reference).

A frequently adopted method is to construct equalized or normalized cDNA libraries (Ko, *Nucleic Acids Res.* 18:5705-5711 (1990), the entirety of which is herein incorporated by reference; Patanjali *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 88:1943-1947 (1991), the entirety of which is herein incorporated by reference). Typically, the cDNA population is normalized by

subtractive hybridization (Schmid *et al.*, *J. Neurochem.* 48:307-312 (1987), the entirety of which is herein incorporated by reference; Fargnoli *et al.*, *Anal. Biochem.* 187:364-373 (1990), the entirety of which is herein incorporated by reference; Travis *et al.*, *Proc. Natl. Acad. Sci (U.S.A.)* 85:1696-1700 (1988), the entirety of which is herein incorporated by reference; Kato, *Eur. J. Neurosci.* 2:704-711 (1990); and Schweinfest *et al.*, *Genet. Anal. Tech. Appl.* 7:64-70 (1990), the entirety of which is herein incorporated by reference). Subtraction represents another method for reducing the population of certain sequences in the cDNA library (Swaroop *et al.*, *Nucleic Acids Res.* 19:1954 (1991), the entirety of which is herein incorporated by reference).

ESTs can be sequenced by a number of methods. Two basic methods may be used for DNA sequencing, the chain termination method of Sanger *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 74:5463-5467 (1977), the entirety of which is herein incorporated by reference and the chemical degradation method of Maxam and Gilbert, *Proc. Nat. Acad. Sci. (U.S.A.)* 74:560-564 (1977), the entirety of which is herein incorporated by reference. Automation and advances in technology such as the replacement of radioisotopes with fluorescence-based sequencing have reduced the effort required to sequence DNA (Craxton, *Methods* 2:20-26 (1991), the entirety of which is herein incorporated by reference; Ju *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 92:4347-4351 (1995), the entirety of which is herein incorporated by reference; Tabor and Richardson, *Proc. Natl. Acad. Sci. (U.S.A.)* 92:6339-6343 (1995), the entirety of which is herein incorporated by reference). Automated sequencers are available from, for example, Pharmacia Biotech, Inc., Piscataway, New Jersey (Pharmacia ALF), LI-COR, Inc., Lincoln, Nebraska (LI-COR 4,000) and Millipore, Bedford, Massachusetts (Millipore BaseStation).

In addition, advances in capillary gel electrophoresis have also reduced the effort required to sequence DNA and such advances provide a rapid high resolution approach for sequencing

DNA samples (Swerdlow and Gesteland, *Nucleic Acids Res.* 18:1415-1419 (1990); Smith, *Nature* 349:812-813 (1991); Luckey *et al.*, *Methods Enzymol.* 218:154-172 (1993); Lu *et al.*, *J. Chromatog. A.* 680:497-501 (1994); Carson *et al.*, *Anal. Chem.* 65:3219-3226 (1993); Huang *et al.*, *Anal. Chem.* 64:2149-2154 (1992); Kheterpal *et al.*, *Electrophoresis* 17:1852-1859 (1996); Quesada and Zhang, *Electrophoresis* 17:1841-1851 (1996); Baba, *Yakugaku Zasshi* 117:265-281 (1997), all of which are herein incorporated by reference in their entirety).

ESTs longer than 150 nucleotides have been found to be useful for similarity searches and mapping (Adams *et al.*, *Science* 252:1651-1656 (1991), herein incorporated by reference). ESTs, which can represent copies of up to the full length transcript, may be partially or completely sequenced. Between 150-450 nucleotides of sequence information is usually generated as this is the length of sequence information that is routinely and reliably produced using single run sequence data. Typically, only single run sequence data is obtained from the cDNA library (Adams *et al.*, *Science* 252:1651-1656 (1991). Automated single run sequencing typically results in an approximately 2-3% error or base ambiguity rate (Boguski *et al.*, *Nature Genetics* 4:332-333 (1993), the entirety of which is herein incorporated by reference).

EST databases have been constructed or partially constructed from, for example, *C. elegans* (McCombie *et al.*, *Nature Genetics* 1:124-131 (1992)), human liver cell line HepG2 (Okubo *et al.*, *Nature Genetics* 2:173-179 (1992)), human brain RNA (Adams *et al.*, *Science* 252:1651-1656 (1991); Adams *et al.*, *Nature* 355:632-635 (1992)), *Arabidopsis*, (Newman *et al.*, *Plant Physiol.* 106:1241-1255 (1994)); and rice (Kurata *et al.*, *Nature Genetics* 8:365-372 (1994)).

### III. SEQUENCE COMPARISONS

A characteristic feature of a DNA sequence is that it can be compared with other DNA sequences. Sequence comparisons can be undertaken by determining the similarity of the test or query sequence with sequences in publicly available or proprietary databases ("similarity analysis") or by searching for certain motifs ("intrinsic sequence analysis")(e.g. *cis* elements)(Coulson, *Trends in Biotechnology* 12:76-80 (1994), the entirety of which is herein incorporated by reference); Birren *et al.*, *Genome Analysis 1*: Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York 543-559 (1997), the entirety of which is herein incorporated by reference).

Similarity analysis includes database search and alignment. Examples of public databases include the DNA Database of Japan (DDBJ)(<http://www.ddbj.nig.ac.jp/>); Genebank (<http://www.ncbi.nlm.nih.gov/Web/Search/Index.html>); and the European Molecular Biology Laboratory Nucleic Acid Sequence Database (EMBL) ([http://www.ebi.ac.uk/ebi\\_docs/embl\\_db/embl-db.html](http://www.ebi.ac.uk/ebi_docs/embl_db/embl-db.html)). Other appropriate databases include dbEST (<http://www.ncbi.nlm.nih.gov/dbEST/index.html>), SwissProt ([http://www.ebi.ac.uk/ebi\\_docs/swisprot\\_db/swisshome.html](http://www.ebi.ac.uk/ebi_docs/swisprot_db/swisshome.html)), PIR (<http://www-nbrt.georgetown.edu/pir/>) and The Institute for Genome Research (<http://www.tigr.org/tdb/tdb.html>)

A number of different search algorithms have been developed, one example of which are the suite of programs referred to as BLAST programs. There are five implementations of BLAST, three designed for nucleotide sequences queries (BLASTN, BLASTX and TBLASTX) and two designed for protein sequence queries (BLASTP and TBLASTN) (Coulson, *Trends in Biotechnology* 12:76-80 (1994); Birren *et al.*, *Genome Analysis 1*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York 543-559 (1997)).

BLASTN takes a nucleotide sequence (the query sequence) and its reverse complement and searches them against a nucleotide sequence database. BLASTN was designed for speed, not maximum sensitivity and may not find distantly related coding sequences. BLASTX takes a nucleotide sequence, translates it in three forward reading frames and three reverse complement reading frames and then compares the six translations against a protein sequence database. BLASTX is useful for sensitive analysis of preliminary (single-pass) sequence data and is tolerant of sequencing errors (Gish and States, *Nature Genetics* 3:266-272 (1993), the entirety of which is herein incorporated by reference). BLASTN and BLASTX may be used in concert for analyzing EST data (Coulson, *Trends in Biotechnology* 12:76-80 (1994); Birren *et al.*, *Genome Analysis* 1:543-559 (1997)).

Given a coding nucleotide sequence and the protein it encodes, it is often preferable to use the protein as the query sequence to search a database because of the greatly increased sensitivity to detect more subtle relationships. This is due to the larger alphabet of proteins (20 amino acids) compared with the alphabet of nucleic acid sequences (4 bases), where it is far easier to obtain a match by chance. In addition, with nucleotide alignments, only a match (positive score) or a mismatch (negative score) is obtained, but with proteins, the presence of conservative amino acid substitutions can be taken into account. Here, a mismatch may yield a positive score if the non-identical residue has physical/chemical properties similar to the one it replaced. Various scoring matrices are used to supply the substitution scores of all possible amino acid pairs. A general purpose scoring system is the BLOSUM62 matrix (Henikoff and Henikoff, *Proteins* 17:49-61 (1993), the entirety of which is herein incorporated by reference), which is currently the default choice for BLAST programs. BLOSUM62 is tailored for alignments of moderately diverged sequences and thus may not yield the best results under all



conditions. Altschul, *J. Mol. Biol.* 36:290-300 (1993), the entirety of which is herein incorporated by reference, describes a combination of three matrices to cover all contingencies. This may improve sensitivity, but at the expense of slower searches. In practice, a single BLOSUM62 matrix is often used but others (PAM40 and PAM250) may be attempted when additional analysis is necessary. Low PAM matrices are directed at detecting very strong but localized sequence similarities, whereas high PAM matrices are directed at detecting long but weak alignments between very distantly related sequences.

Homologues in other organisms are available that can be used for comparative sequence analysis. Multiple alignments are performed to study similarities and differences in a group of related sequences. CLUSTAL W is a multiple sequence alignment package that performs progressive multiple sequence alignments based on the method of Feng and Doolittle, *J. Mol. Evol.* 25:351-360 (1987), the entirety of which is herein incorporated by reference. Each pair of sequences is aligned and the distance between each pair is calculated; from this distance matrix, a guide tree is calculated and all of the sequences are progressively aligned based on this tree. A feature of the program is its sensitivity to the effect of gaps on the alignment; gap penalties are varied to encourage the insertion of gaps in probable loop regions instead of in the middle of structured regions. Users can specify gap penalties, choose between a number of scoring matrices, or supply their own scoring matrix for both pairwise alignments and multiple alignments. CLUSTAL W for UNIX and VMS systems is available at: [ftp.ebi.ac.uk](http://ftp.ebi.ac.uk). Another program is MACAW (Schuler *et al.*, *Proteins Struct. Func. Genet.* 9:180-190 (1991), the entirety of which is herein incorporated by reference, for which both Macintosh and Microsoft Windows versions are available. MACAW uses a graphical interface, provides a choice of several

alignment algorithms and is available by anonymous ftp at: [ncbi.nlm.nih.gov](ftp://ncbi.nlm.nih.gov)  
(directory/pub/macaw).

Sequence motifs are derived from multiple alignments and can be used to examine individual sequences or an entire database for subtle patterns. With motifs, it is sometimes possible to detect distant relationships that may not be demonstrable based on comparisons of primary sequences alone. Currently, the largest collection of sequence motifs in the world is PROSITE (Bairoch and Bucher, *Nucleic Acid Research* 22:3583-3589 (1994), the entirety of which is herein incorporated by reference). PROSITE may be accessed via either the ExPASy server on the World Wide Web or anonymous ftp site. Many commercial sequence analysis packages also provide search programs that use PROSITE data.

A resource for searching protein motifs is the BLOCKS E-mail server developed by Henikoff, *Trends Biochem Sci.* 18:267-268 (1993), the entirety of which is herein incorporated by reference; Henikoff and Henikoff, *Nucleic Acid Research* 19:6565-6572 (1991), the entirety of which is herein incorporated by reference; Henikoff and Henikoff, *Proteins* 17:49-61 (1993). BLOCKS searches a protein or nucleotide sequence against a database of protein motifs or "blocks." Blocks are defined as short, ungapped multiple alignments that represent highly conserved protein patterns. The blocks themselves are derived from entries in PROSITE as well as other sources. Either a protein query or a nucleotide query can be submitted to the BLOCKS server; if a nucleotide sequence is submitted, the sequence is translated in all six reading frames and motifs are sought for these conceptual translations. Once the search is completed, the server will return a ranked list of significant matches, along with an alignment of the query sequence to the matched BLOCKS entries.

Conserved protein domains can be represented by two-dimensional matrices, which measure either the frequency or probability of the occurrences of each amino acid residue and deletions or insertions in each position of the domain. This type of model, when used to search against protein databases, is sensitive and usually yields more accurate results than simple motif searches. Two popular implementations of this approach are profile searches such as GCG program ProfileSearch and Hidden Markov Models (HMMs)(Krough *et al.*, *J. Mol. Biol.* 235:1501-1531, (1994); Eddy, *Current Opinion in Structural Biology* 6:361-365, (1996), both of which are herein incorporated by reference in their entirety). In both cases, a large number of common protein domains have been converted into profiles, as present in the PROSITE library, or HMM models, as in the Pfam protein domain library (Sonnhammer *et al.*, *Proteins* 28:405-420 (1997), the entirety of which is herein incorporated by reference). Pfam contains more than 500 HMM models for enzymes, transcription factors, signal transduction molecules and structural proteins. Protein databases can be queried with these profiles or HMM models, which will identify proteins containing the domain of interest. For example, HMMSW or HMMFS, two programs in a public domain package called HMMER (Sonnhammer *et al.*, *Proteins* 28:405-420 (1997)) can be used.

PROSITE and BLOCKS represent collected families of protein motifs. Thus, searching these databases entails submitting a single sequence to determine whether or not that sequence is similar to the members of an established family. Programs working in the opposite direction compare a collection of sequences with individual entries in the protein databases. An example of such a program is the Motif Search Tool, or MoST (Tatusov *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 91:12091-12095 (1994), the entirety of which is herein incorporated by reference). On the basis of an aligned set of input sequences, a weight matrix is calculated by using one of four

methods (selected by the user). A weight matrix is simply a representation, position by position of how likely a particular amino acid will appear. The calculated weight matrix is then used to search the databases. To increase sensitivity, newly found sequences are added to the original data set, the weight matrix is recalculated and the search is performed again. This procedure continues until no new sequences are found.

### **SUMMARY OF THE INVENTION**

The present invention provides a substantially purified nucleic acid molecule that encodes a maize or a soybean enzyme or fragment thereof, wherein the maize or the soybean enzyme is selected from the group consisting of: (a) triose phosphate isomerase; (b) fructose 1,6-bisphosphate aldolase; (c) fructose 1,6-bisphosphate; (d) fructose 6-phosphate 2-kinase; (e) phosphoglucosomerase; (f) vacuolar H<sup>+</sup> translocating-pyrophosphatase; (g) pyrophosphate-dependent fructose-6-phosphate phosphotransferase; (h) invertase; (i) sucrose synthase; (j) hexokinase; (k) fructokinase; (l) NDP-kinase; (m) glucose-6-phosphate 1-dehydrogenase; (n) phosphoglucomutase and (o) UDP-glucose pyrophosphorylase.

The present invention also provides a substantially purified nucleic acid molecule that encodes a plant sucrose pathway enzyme or fragment thereof, wherein the nucleic acid molecule is selected from the group consisting of a nucleic acid molecule that encodes a maize or a soybean triose phosphate isomerase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate aldolase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 6-phosphate 2-kinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean phosphoglucosomerase enzyme or fragment thereof, a nucleic acid molecule that encodes a

maize or a soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean invertase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean sucrose synthase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean hexokinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructokinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean NDP-kinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean glucose-6-phosphate 1-dehydrogenase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean phosphoglucomutase enzyme or fragment thereof and a nucleic acid molecule that encodes a maize or a soybean UDP-glucose pyrophosphorylase enzyme or fragment thereof.

The present invention also provides a substantially purified maize or soybean enzyme or fragment thereof, wherein the maize or soybean enzyme is selected from the group consisting of (a) triose phosphate isomerase; (b) fructose 1,6-bisphosphate aldolase; (c) fructose 1,6-bisphosphate; (d) fructose 6-phosphate 2-kinase; (e) phosphoglucoisomerase; (f) vacuolar H<sup>+</sup> translocating-pyrophosphatase; (g) pyrophosphate-dependent fructose-6-phosphate phosphotransferase; (h) invertase; (i) sucrose synthase; (j) hexokinase; (k) fructokinase; (l) NDP-kinase; (m) glucose-6-phosphate 1-dehydrogenase; (n) phosphoglucomutase and (o) UDP-glucose pyrophosphorylase.

The present invention also provides a substantially purified maize or soybean sucrose pathway protein or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic

acid sequence selected from the group consisting of a complement of SEQ ID NO: 1 through SEQ ID NO: 2814.

The present invention also provides a substantially purified maize or soybean triose phosphate isomerase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1 through SEQ ID NO: 206 and SEQ ID NO: 1538 through SEQ ID NO: 1707.

The present invention also provides a substantially purified maize or soybean triose phosphate isomerase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 206 and SEQ ID NO: 1538 through SEQ ID NO: 1707.

The present invention also provides a substantially purified maize or soybean fructose 1,6-bisphosphate aldolase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 207 through SEQ ID NO: 232 and SEQ ID NO: 1708 through SEQ ID NO: 2113.

The present invention also provides a substantially purified maize or soybean fructose 1,6-bisphosphate aldolase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting SEQ ID NO: 207 through SEQ ID NO: 232 and SEQ ID NO: 1708 through SEQ ID NO: 2113.

The present invention also provides a substantially purified maize or soybean fructose 1,6-bisphosphate enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule

having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 233 through SEQ ID NO: 258 and SEQ ID NO: 2114 through SEQ ID NO: 2162.

The present invention also provides a substantially purified maize or soybean fructose 1,6-bisphosphate e enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 233 through SEQ ID NO: 258 and SEQ ID NO: 2114 through SEQ ID NO: 2162.

The present invention also provides a substantially purified maize or soybean fructose 6-phosphate 2-kinase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 259 through SEQ ID NO: 275 and SEQ ID NO: 2163 through SEQ ID NO: 2166.

The present invention also provides a substantially purified maize or soybean fructose 6-phosphate 2-kinase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 259 through SEQ ID NO: 275 and SEQ ID NO: 2163 through SEQ ID NO: 2166.

The present invention also provides a substantially purified maize or soybean phosphoglucoisomerase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 276 through SEQ ID NO: 340 and SEQ ID NO: 2167 through SEQ ID NO: 2182.

The present invention also provides a substantially purified maize or soybean phosphoglucoisomerase enzyme or fragment thereof encoded by a nucleic acid sequence selected

from the group consisting of SEQ ID NO: 276 through SEQ ID NO: 340 and SEQ ID NO: 2167 through SEQ ID NO: 2182.

The present invention also provides a substantially purified maize or soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 341 through SEQ ID NO: 497 and SEQ ID NO: 2183 through SEQ ID NO: 2241.

The present invention also provides a substantially purified maize or soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 341 through SEQ ID NO: 497 and SEQ ID NO: 2183 through SEQ ID NO: 2241.

The present invention also provides a substantially purified maize or soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 498 through SEQ ID NO: 507 and SEQ ID NO: 2442.

The present invention also provides a substantially purified maize or soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 498 through SEQ ID NO: 507 and SEQ ID NO: 2442.



The present invention also provides a substantially purified maize or soybean invertase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 508 through SEQ ID NO: 510 and SEQ ID NO: 2243 through SEQ ID NO: 2254.

The present invention also provides a substantially purified maize or soybean invertase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 508 through SEQ ID NO: 510 and SEQ ID NO: 2243 through SEQ ID NO: 2254.

The present invention also provides a substantially purified maize or soybean sucrose synthase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 511 through SEQ ID NO: 1086 and SEQ ID NO: 2255 through SEQ ID NO: 2590.

The present invention also provides a substantially purified maize or soybean sucrose synthase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 511 through SEQ ID NO: 1086 and SEQ ID NO: 2255 through SEQ ID NO: 2590.

The present invention also provides a substantially purified maize or soybean hexokinase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1087 through SEQ ID NO: 1135 and SEQ ID NO: 2591 through SEQ ID NO: 2634.

The present invention also provides a substantially purified maize or soybean hexokinase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1087 through SEQ ID NO: 1135 and SEQ ID NO: 2591 through SEQ ID NO: 2634.

The present invention also provides a substantially purified maize or soybean fructokinase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1136 through SEQ ID NO: 1215 and SEQ ID NO: 2635 through SEQ ID NO: 2678.

The present invention also provides a substantially purified maize or soybean fructokinase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1136 through SEQ ID NO: 1215 and SEQ ID NO: 2635 through SEQ ID NO: 2678.

The present invention also provides a substantially purified maize or soybean NDP-kinase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1216 through SEQ ID NO: 1251 and SEQ ID NO: 2679 through SEQ ID NO: 2681.

The present invention also provides a substantially purified maize or soybean NDP-kinase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1216 through SEQ ID NO: 1251 and SEQ ID NO: 2679 through SEQ ID NO: 2681.

The present invention also provides a substantially purified maize or soybean glucose-6-phosphate 1-dehydrogenase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1252 through SEQ ID NO: 1254 and SEQ ID NO: 2682 through SEQ ID NO: 2689.

The present invention also provides a substantially purified maize or soybean glucose-6-phosphate 1-dehydrogenase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1252 through SEQ ID NO: 1254 and SEQ ID NO: 2682 through SEQ ID NO: 2689.

The present invention also provides a substantially purified maize or soybean phosphoglucomutase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1255 through SEQ ID NO: 1360 and SEQ ID NO: 2690 through SEQ ID NO: 2740.

The present invention also provides a substantially purified maize or soybean phosphoglucomutase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1255 through SEQ ID NO: 1360 and SEQ ID NO: 2690 through SEQ ID NO: 2740.

The present invention also provides a substantially purified maize or soybean UDP-glucose pyrophosphorylase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of

SEQ ID NO: 1361 through SEQ ID NO: 1537 and SEQ ID NO: 2741 through SEQ ID NO: 2814.

The present invention also provides a substantially purified maize or soybean UDP-glucose pyrophosphorylase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1361 through SEQ ID NO: 1537 and SEQ ID NO: 2741 through SEQ ID NO: 2814.

The present invention also provides a purified antibody or fragment thereof which is capable of specifically binding to a maize or soybean enzyme or fragment thereof, wherein the maize or soybean enzyme or fragment thereof is encoded by a nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean triose phosphate isomerase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1 through SEQ ID NO: 206 and SEQ ID NO: 1538 through SEQ ID NO: 1707 and a maize or soybean triose phosphate isomerase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 206 and SEQ ID NO: 1538 through SEQ ID NO: 1707.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean fructose 1,6-bisphosphate aldolase enzyme or fragment thereof encoded by a first nucleic acid molecule

which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 207 through SEQ ID NO: 232 and SEQ ID NO: 1708 through SEQ ID NO: 2113 and a maize or soybean fructose 1,6-bisphosphate aldolase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 207 through SEQ ID NO: 232 and SEQ ID NO: 1708 through SEQ ID NO: 2113.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean fructose 1,6-bisphosphate enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 233 through SEQ ID NO: 258 and SEQ ID NO: 2114 through SEQ ID NO: 2162 and a maize or soybean fructose 1,6-bisphosphate enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 233 through SEQ ID NO: 258 and SEQ ID NO: 2114 through SEQ ID NO: 2162.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean fructose 6-phosphate 2-kinase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 259 through SEQ ID NO: 275 and SEQ ID NO: 2163 through SEQ ID NO: 2166 and a maize or soybean fructose 6-phosphate 2-kinase enzyme or fragment thereof encoded by a

nucleic acid sequence selected from the group consisting of SEQ ID NO: 259 through SEQ ID NO: 275 and SEQ ID NO: 2163 through SEQ ID NO: 2166.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean phosphoglucosyltransferase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 276 through SEQ ID NO: 340 and SEQ ID NO: 2167 through SEQ ID NO: 2182 and a maize or soybean phosphoglucosyltransferase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 276 through SEQ ID NO: 340 and SEQ ID NO: 2167 through SEQ ID NO: 2182.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 341 through SEQ ID NO: 497 and SEQ ID NO: 2183 through SEQ ID NO: 2241 and a maize or soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 341 through SEQ ID NO: 497 and SEQ ID NO: 2183 through SEQ ID NO: 2241.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or fragment thereof

encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 498 through SEQ ID NO: 507 and SEQ ID NO: 2442 and a maize or soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 498 through SEQ ID NO: 507 and SEQ ID NO: 2442.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean invertase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 508 through SEQ ID NO: 510 and SEQ ID NO: 2243 through SEQ ID NO: 2254 and a maize or soybean invertase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 508 through SEQ ID NO: 510 and SEQ ID NO: 2243 through SEQ ID NO: 2254.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean sucrose synthase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 511 through SEQ ID NO: 1086 and SEQ ID NO: 2255 through SEQ ID NO: 2590 and a maize or soybean sucrose synthase enzyme or fragment thereof encoded by a nucleic acid sequence selected from

the group consisting of SEQ ID NO: 511 through SEQ ID NO: 1086 and SEQ ID NO: 2255 through SEQ ID NO: 2590.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean hexokinase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1087 through SEQ ID NO: 1135 and SEQ ID NO: 2591 through SEQ ID NO: 2634 and a maize or soybean hexokinase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1087 through SEQ ID NO: 1135 and SEQ ID NO: 2591 through SEQ ID NO: 2634.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean fructokinase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1136 through SEQ ID NO: 1215 and SEQ ID NO: 2635 through SEQ ID NO: 2678 and a maize or soybean fructokinase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1136 through SEQ ID NO: 1215 and SEQ ID NO: 2635 through SEQ ID NO: 2678.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean NDP-kinase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically



hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1216 through SEQ ID NO: 1251 and SEQ ID NO: 2679 through SEQ ID NO: 2681 and a maize or soybean NDP-kinase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1216 through SEQ ID NO: 1251 and SEQ ID NO: 2679 through SEQ ID NO: 2681.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean glucose-6-phosphate 1-dehydrogenase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1252 through SEQ ID NO: 1254 and SEQ ID NO: 2682 through SEQ ID NO: 2689 and a maize or soybean glucose-6-phosphate 1-dehydrogenase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1252 through SEQ ID NO: 1254 and SEQ ID NO: 2682 through SEQ ID NO: 2689.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean phosphoglucomutase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1255 through SEQ ID NO: 1360 and SEQ ID NO: 2690 through SEQ ID NO: 2740 and a maize or soybean phosphoglucomutase enzyme or fragment thereof encoded by a nucleic acid

sequence selected from the group consisting of SEQ ID NO: 1255 through SEQ ID NO: 1360 and SEQ ID NO: 2690 through SEQ ID NO: 2740.

The present invention also provides a substantially purified antibody or fragment thereof, the antibody or fragment thereof capable of specifically binding to a maize or a soybean UDP-glucose pyrophosphorylase enzyme or fragment thereof encoded by a first nucleic acid molecule which specifically hybridizes to a second nucleic acid molecule, the second nucleic acid molecule having a nucleic acid sequence selected from the group consisting of a complement of SEQ ID NO: 1361 through SEQ ID NO: 1537 and SEQ ID NO: 2741 through SEQ ID NO: 2814 and a maize or soybean UDP-glucose pyrophosphorylase enzyme or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1361 through SEQ ID NO: 1537 and SEQ ID NO: 2741 through SEQ ID NO: 2814.

The present invention also provides a transformed plant having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule; (B) a structural nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of (a) a nucleic acid sequence which encodes for triose phosphate isomerase or fragment thereof; (b) a nucleic acid sequence which encodes for fructose 1,6-bisphosphate aldolase or fragment thereof; (c) a nucleic acid sequence which encodes for fructose 1,6-bisphosphate or fragment thereof; (d) a nucleic acid sequence which encodes for fructose 6-phosphate 2-kinase or fragment thereof; (e) a nucleic acid sequence which encodes for phosphoglucoisomerase or fragment thereof; (f) a nucleic acid sequence which encodes for vacuolar H<sup>+</sup> translocating-pyrophosphatase or fragment thereof; (g) a nucleic acid sequence which encodes for pyrophosphate-dependent fructose-6-phosphate phosphotransferase or fragment thereof; (h) a nucleic acid sequence which encodes for invertase or fragment thereof;

(i) a nucleic acid sequence which encodes for sucrose synthase or fragment thereof; (j) a nucleic acid sequence which encodes for hexokinase or fragment thereof; (k) a nucleic acid sequence which encodes for fructokinase or fragment thereof; (l) a nucleic acid sequence which encodes for NDP-kinase or fragment thereof; (m) a nucleic acid sequence which encodes for glucose-6-phosphate 1-dehydrogenase or fragment thereof; (n) a nucleic acid sequence which encodes for phosphoglucomutase or fragment thereof (o) a nucleic acid sequence which encodes for UDP-glucose pyrophosphorylase or fragment thereof and (p) a nucleic acid sequence which is complementary to any of the nucleic acid sequences of (a) through (o); and (C) a 3' non-translated sequence that functions in the plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

The present invention also provides a transformed plant having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule; which is linked to (B) a structural nucleic acid molecule, wherein the structural nucleic acid molecule encodes a plant sucrose pathway enzyme or fragment thereof, the structural nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814 or fragment thereof; which is linked to (C) a 3' non-translated sequence that functions in the plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

The present invention also provides a transformed plant having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule; which is linked to (B) a structural nucleic acid molecule, wherein the structural nucleic acid molecule is selected from the group consisting of a nucleic

acid molecule that encodes a maize or a soybean triose phosphate isomerase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate aldolase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 6-phosphate 2-kinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean phosphoglucosomerase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean invertase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean sucrose synthase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean hexokinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructokinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean NDP-kinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean glucose-6-phosphate 1-dehydrogenase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean phosphoglucomutase enzyme or fragment thereof and a nucleic acid molecule that encodes a maize or a soybean UDP-glucose pyrophosphorylase enzyme or fragment thereof; which is linked to (C) a 3' non-translated sequence that functions in the plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

The present invention also provides a transformed plant having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions in a plant cell to cause the

production of a mRNA molecule; which is linked to (B) a transcribed nucleic acid molecule with a transcribed strand and a non-transcribed strand, wherein the transcribed strand is complementary to a nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814 or fragment thereof; which is linked to (C) a 3' non-translated sequence that functions in plant cells to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

The present invention also provides a transformed plant having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule; which is linked to: (B) a transcribed nucleic acid molecule with a transcribed strand and a non-transcribed strand, wherein a transcribed mRNA of the transcribed strand is complementary to an endogenous mRNA molecule having a nucleic acid sequence selected from the group consisting of an endogenous mRNA molecule that encodes a maize or a soybean triose phosphate isomerase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean fructose 1,6-bisphosphate aldolase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean fructose 1,6-bisphosphate enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean fructose 6-phosphate 2-kinase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean phosphoglucisomerase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean invertase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a

soybean sucrose synthase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean hexokinase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean fructokinase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean NDP-kinase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean glucose-6-phosphate 1-dehydrogenase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean phosphoglucomutase enzyme or fragment thereof and an endogenous mRNA molecule that encodes a maize or a soybean UDP-glucose pyrophosphorylase enzyme or fragment thereof; which is linked to (C) a 3' non-translated sequence that functions in the plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

The present invention also provides a method for determining a level or pattern in a plant cell of an enzyme in a plant metabolic pathway comprising: (A) incubating, under conditions permitting nucleic acid hybridization, a marker nucleic acid molecule, the marker nucleic acid molecule selected from the group of marker nucleic acid molecules which specifically hybridize to a nucleic acid molecule having the nucleic acid sequence of SEQ ID NO: 1 through SEQ ID NO: 2814 or compliments thereof, with a complementary nucleic acid molecule obtained from the plant cell or plant tissue, wherein nucleic acid hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant cell or plant tissue permits the detection of an mRNA for the enzyme; (B) permitting hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant cell or plant tissue; and (C) detecting the level or pattern of the complementary nucleic

acid, wherein the detection of the complementary nucleic acid is predictive of the level or pattern of the enzyme in the plant metabolic pathway.

The present invention also provides a method for determining a level or pattern of a plant sucrose pathway enzyme in a plant cell or plant tissue comprising: (A) incubating, under conditions permitting nucleic acid hybridization, a marker nucleic acid molecule, the marker nucleic acid molecule having a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof or fragment of either, with a complementary nucleic acid molecule obtained from the plant cell or plant tissue, wherein nucleic acid hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant cell or plant tissue permits the detection of the plant sucrose pathway enzyme; (B) permitting hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant cell or plant tissue; and (C) detecting the level or pattern of the complementary nucleic acid, wherein the detection of the complementary nucleic acid is predictive of the level or pattern of the plant sucrose pathway enzyme.

The present invention also provides a method for determining a level or pattern of a plant sucrose pathway enzyme in a plant cell or plant tissue comprising: (A) incubating, under conditions permitting nucleic acid hybridization, a marker nucleic acid molecule, the marker nucleic acid molecule comprising a nucleic acid molecule that encodes a maize or a soybean triose phosphate isomerase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate aldolase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate enzyme or complement thereof or fragment of either, a

nucleic acid molecule that encodes a maize or a soybean fructose 6-phosphate 2-kinase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean phosphoglucosomerase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean invertase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean sucrose synthase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean hexokinase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean fructokinase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean NDP-kinase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean glucose-6-phosphate 1-dehydrogenase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean phosphoglucomutase enzyme or complement thereof or fragment of either and a nucleic acid molecule that encodes a maize or a soybean UDP-glucose pyrophosphorylase enzyme or complement thereof or fragment of either, with a complementary nucleic acid molecule obtained from the plant cell or plant tissue, wherein nucleic acid hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant cell or plant tissue permits the detection of the plant sucrose pathway enzyme; (B) permitting hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant cell or plant



tissue; and (C) detecting the level or pattern of the complementary nucleic acid, wherein the detection of the complementary nucleic acid is predictive of the level or pattern of the plant sucrose pathway enzyme.

The present invention also provides a method for determining a level or pattern of a plant sucrose pathway enzyme in a plant cell or plant tissue under evaluation which comprises assaying the concentration of a molecule, whose concentration is dependent upon the expression of a gene, the gene specifically hybridizes to a nucleic acid molecule having a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof, in comparison to the concentration of that molecule present in a reference plant cell or a reference plant tissue with a known level or pattern of the plant sucrose pathway enzyme, wherein the assayed concentration of the molecule is compared to the assayed concentration of the molecule in the reference plant cell or reference plant tissue with the known level or pattern of the plant sucrose pathway enzyme.

The present invention also provides a method for determining a level or pattern of a plant sucrose pathway enzyme in a plant cell or plant tissue under evaluation which comprises assaying the concentration of a molecule, whose concentration is dependent upon the expression of a gene, the gene specifically hybridizes to a nucleic acid molecule selected from the group consisting of a nucleic acid molecule that encodes a maize or a soybean triose phosphate isomerase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate aldolase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 6-phosphate 2-kinase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean

phosphoglucisomerase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean invertase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean sucrose synthase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean hexokinase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean fructokinase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean NDP-kinase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean glucose-6-phosphate 1-dehydrogenase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean phosphoglucumutase enzyme or complement thereof and a nucleic acid molecule that encodes a maize or a soybean UDP-glucose pyrophosphorylase enzyme or complement thereof, in comparison to the concentration of that molecule present in a reference plant cell or a reference plant tissue with a known level or pattern of the plant sucrose pathway enzyme, wherein the assayed concentration of the molecule is compared to the assayed concentration of the molecule in the reference plant cell or the reference plant tissue with the known level or pattern of the plant sucrose pathway enzyme.

The present invention provides a method of determining a mutation in a plant whose presence is predictive of a mutation affecting a level or pattern of a protein comprising the steps: (A) incubating, under conditions permitting nucleic acid hybridization, a marker nucleic acid, the marker nucleic acid selected from the group of marker nucleic acid molecules which specifically hybridize to a nucleic acid molecule having a nucleic acid sequence selected from the group of

SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof and a complementary nucleic acid molecule obtained from the plant, wherein nucleic acid hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant permits the detection of a polymorphism whose presence is predictive of a mutation affecting the level or pattern of the plant sucrose pathway enzyme in the plant; (B) permitting hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant; and (C) detecting the presence of the polymorphism, wherein the detection of the polymorphism is predictive of the mutation.

The present invention also provides a method for determining a mutation in a plant whose presence is predictive of a mutation affecting the level or pattern of a plant sucrose pathway enzyme comprising the steps: (A) incubating, under conditions permitting nucleic acid hybridization, a marker nucleic acid molecule, the marker nucleic acid molecule comprising a nucleic acid molecule that is linked to a gene, the gene specifically hybridizes to a nucleic acid molecule having a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof and a complementary nucleic acid molecule obtained from the plant, wherein nucleic acid hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant permits the detection of a polymorphism whose presence is predictive of a mutation affecting the level or pattern of the plant sucrose pathway enzyme in the plant; (B) permitting hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant; and (C) detecting the presence of the polymorphism, wherein the detection of the polymorphism is predictive of the mutation.

The present invention also provides a method for determining a mutation in a plant whose presence is predictive of a mutation affecting the level or pattern of a plant sucrose pathway enzyme comprising the steps: (A) incubating, under conditions permitting nucleic acid hybridization, a marker nucleic acid molecule, the marker nucleic acid molecule comprising a nucleic acid molecule that is linked to a gene, the gene specifically hybridizes to a nucleic acid molecule selected from the group consisting of a nucleic acid molecule that encodes a maize or a soybean triose phosphate isomerase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate aldolase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 6-phosphate 2-kinase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean phosphoglucosomerase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean invertase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean sucrose synthase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean hexokinase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean fructokinase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean NDP-kinase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean glucose-6-phosphate 1-dehydrogenase enzyme or complement thereof, a nucleic acid molecule that encodes a maize or a soybean phosphoglucomutase enzyme or

complement thereof and a nucleic acid molecule that encodes a maize or a soybean UDP-glucose pyrophosphorylase enzyme or complement thereof and a complementary nucleic acid molecule obtained from the plant, wherein nucleic acid hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant permits the detection of a polymorphism whose presence is predictive of a mutation affecting the level or pattern of the plant sucrose pathway enzyme in the plant; (B) permitting hybridization between the marker nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant; and (C) detecting the presence of the polymorphism, wherein the detection of the polymorphism is predictive of the mutation.

The present invention also provides a method of producing a plant containing an overexpressed protein comprising: (A) transforming the plant with a functional nucleic acid molecule, wherein the functional nucleic acid molecule comprises a promoter region, wherein the promoter region is linked to a structural region, wherein the structural region has a nucleic acid sequence selected from group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814 wherein the structural region is linked to a 3' non-translated sequence that functions in the plant to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of a mRNA molecule; and wherein the functional nucleic acid molecule results in overexpression of the protein; and (B) growing the transformed plant.

The present invention also provides a method of producing a plant containing an overexpressed plant sucrose enzyme comprising: (A) transforming the plant with a functional nucleic acid molecule, wherein the functional nucleic acid molecule comprises a promoter region, wherein the promoter region is linked to a structural region, wherein the structural region comprises a nucleic acid molecule having a nucleic acid sequence selected from the group

consisting of SEQ ID NO: 1 through SEQ ID NO: 2814 or fragment thereof; wherein the structural region is linked to a 3' non-translated sequence that functions in the plant to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of a mRNA molecule; and wherein the functional nucleic acid molecule results in overexpression of the plant sucrose pathway enzyme; and (B) growing the transformed plant.

The present invention also provides a method of producing a plant containing an overexpressed plant sucrose pathway enzyme comprising: (A) transforming the plant with a functional nucleic acid molecule, wherein the functional nucleic acid molecule comprises a promoter region, wherein the promoter region is linked to a structural region, wherein the structural region comprises a nucleic acid molecule selected from the group consisting of a nucleic acid molecule that encodes a maize or a soybean triose phosphate isomerase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate aldolase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 6-phosphate 2-kinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean phosphoglucoisomerase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean invertase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean sucrose synthase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean hexokinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a

soybean fructokinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean NDP-kinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean glucose-6-phosphate 1-dehydrogenase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean phosphoglucomutase enzyme or fragment thereof and a nucleic acid molecule that encodes a maize or a soybean UDP-glucose pyrophosphorylase enzyme or fragment thereof, wherein the structural region is linked to a 3' non-translated sequence that functions in the plant to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of a mRNA molecule; and wherein the functional nucleic acid molecule results in overexpression of the plant sucrose pathway enzyme protein; and (B) growing the transformed plant.

The present invention also provides a method of producing a plant containing reduced levels of a plant sucrose pathway enzyme comprising: (A) transforming the plant with a functional nucleic acid molecule, wherein the functional nucleic acid molecule comprises a promoter region, wherein the promoter region is linked to a structural region, wherein the structural region comprises a nucleic acid molecule having a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814; wherein the structural region is linked to a 3' non-translated sequence that functions in the plant to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of a mRNA molecule; and wherein the functional nucleic acid molecule results in co-suppression of the plant sucrose pathway enzyme protein; and (B) growing the transformed plant.

The present invention also provides a method of producing a plant containing reduced levels of a plant sucrose pathway enzyme comprising: (A) transforming the plant with a functional nucleic acid molecule, wherein the functional nucleic acid molecule comprises a

[illegible]



mRNA molecule; and wherein the functional nucleic acid molecule results in co-suppression of the plant sucrose pathway enzyme; and (B) growing the transformed plant.

The present invention also provides a method for reducing expression of a plant sucrose pathway enzyme in a plant comprising: (A) transforming the plant with a nucleic acid molecule, the nucleic acid molecule having an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule, wherein the exogenous promoter region is linked to a transcribed nucleic acid molecule having a transcribed strand and a non-transcribed strand, wherein the transcribed strand is complementary to a nucleic acid molecule having a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof or fragments of either and the transcribed strand is complementary to an endogenous mRNA molecule; and wherein the transcribed nucleic acid molecule is linked to a 3' non-translated sequence that functions in the plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of a mRNA molecule; and (B) growing the transformed plant.

The present invention also provides a method for reducing expression of a plant sucrose pathway enzyme in a plant comprising: (A) transforming the plant with a nucleic acid molecule, the nucleic acid molecule having an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule, wherein the exogenous promoter region is linked to a transcribed nucleic acid molecule having a transcribed strand and a non-transcribed strand, wherein a transcribed mRNA of the transcribed strand is complementary to a nucleic acid molecule selected from the group consisting of an endogenous mRNA molecule that encodes a maize or a soybean triose phosphate isomerase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean fructose 1,6-bisphosphate aldolase enzyme

or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean fructose 1,6-bisphosphate enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean fructose 6-phosphate 2-kinase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean phosphoglucosomerase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean invertase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean sucrose synthase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean hexokinase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean fructokinase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean NDP-kinase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean glucose-6-phosphate 1-dehydrogenase enzyme or fragment thereof, an endogenous mRNA molecule that encodes a maize or a soybean phosphoglucomutase enzyme or fragment thereof and an endogenous mRNA molecule that encodes a maize or a soybean UDP-glucose pyrophosphorylase enzyme or fragment thereof, and wherein the transcribed nucleic acid molecule is linked to a 3' non-translated sequence that functions in the plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of a mRNA molecule; and (B) growing the transformed plant.

The present invention also provides a method of determining an association between a polymorphism and a plant trait comprising: (A) hybridizing a nucleic acid molecule specific for

the polymorphism to genetic material of a plant, wherein the nucleic acid molecule has a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof or fragment of either; and (B) calculating the degree of association between the polymorphism and the plant trait.

The present invention also provides a method of determining an association between a polymorphism and a plant trait comprising: (A) hybridizing a nucleic acid molecule specific for the polymorphism to genetic material of a plant, wherein the nucleic acid molecule is selected from the group consisting of a nucleic acid molecule that encodes a maize or a soybean triose phosphate isomerase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate aldolase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean fructose 6-phosphate 2-kinase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean phosphoglucosomerase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean invertase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean sucrose synthase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean hexokinase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a

soybean fructokinase enzyme or complement thereof or fragment of either f, a nucleic acid molecule that encodes a maize or a soybean NDP-kinase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean glucose-6-phosphate 1-dehydrogenase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean phosphoglucomutase enzyme or complement thereof or fragment of either and a nucleic acid molecule that encodes a maize or a soybean UDP-glucose pyrophosphorylase enzyme or complement thereof or fragment of either and (B) calculating the degree of association between the polymorphism and the plant trait.

The present invention also provides a method of isolating a nucleic acid that encodes a plant sucrose pathway enzyme or fragment thereof comprising: (A) incubating under conditions permitting nucleic acid hybridization, a first nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof or fragment of either with a complementary second nucleic acid molecule obtained from a plant cell or plant tissue; (B) permitting hybridization between the first nucleic acid molecule and the second nucleic acid molecule obtained from the plant cell or plant tissue; and (C) isolating the second nucleic acid molecule.

The present invention also provides a method of isolating a nucleic acid molecule that encodes a plant sucrose pathway enzyme or fragment thereof comprising: (A) incubating under conditions permitting nucleic acid hybridization, a first nucleic acid molecule selected from the group consisting of a nucleic acid molecule that encodes a maize or a soybean triose phosphate isomerase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate aldolase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-

bisphosphate enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean fructose 6-phosphate 2-kinase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean phosphoglucoisomerase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean vacuolar  $H^+$  translocating-pyrophosphatase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean invertase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean sucrose synthase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean hexokinase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean fructokinase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean NDP-kinase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean glucose-6-phosphate 1-dehydrogenase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean phosphoglucomutase enzyme or complement thereof or fragment of either and a nucleic acid molecule that encodes a maize or a soybean UDP-glucose pyrophosphorylase enzyme or complement thereof or fragment of either, with a complementary second nucleic acid molecule obtained from a plant cell or plant tissue; (B) permitting hybridization between the plant sucrose pathway nucleic acid molecule and the complementary nucleic acid molecule obtained from the plant cell or plant tissue; and (C) isolating the second nucleic acid molecule.

## **DETAILED DESCRIPTION OF THE INVENTION**

### **Definitions and Agents of the Present Invention**

#### **Definitions:**

As used herein, a sucrose pathway enzyme is any enzyme that is associated with the synthesis or degradation of sucrose.

As used herein, a sucrose synthesis enzyme is any enzyme that is associated with the synthesis of sucrose.

As used herein, a sucrose degradation enzyme is any enzyme that is associated with the degradation of sucrose.

As used herein, triose phosphate isomerase is any enzyme that maintains at equilibrium the pool of triose phosphates, dihydroxyacetone phosphate ("DHAP"), and glyceraldehyde-3-phosphate ("GAP") within the cytoplasm.

As used herein, fructose 1,6-bisphosphate aldolase is any enzyme that catalyzes an aldol condensation of DHAP and GAP to form fructose 1,6-bisphosphate ("F1,6BP").

As used herein, fructose-1,6-bisphosphatase ("FBPase") is any enzyme that catalyzes the cleavage of phosphate from the C1 carbon of fructose-1,6-bisphosphate to form fructose-6-phosphate ("F6P").

As used herein, fructose 6-phosphate 2-kinase is any enzyme that controls the concentration of fructose 2,6-bisphosphate.

As used herein, phosphoglucoisomerase is any enzyme that maintains glucose-6-phosphate ("G6P") and glucose-1-phosphate ("G1P") in equilibrium with the F6P pool.

As used herein, vacuolar H<sup>+</sup> translocating-pyrophosphatase is any enzyme that utilizes pyrophosphate to sustain a proton gradient formed within the vacuolar membrane.

As used herein, pyrophosphate-dependent fructose-6-phosphate phosphotransferase is any enzyme that catalyzes the reversible production of F1,6BP and Pi from F6P and PPI.

As used herein, invertase is any enzyme that irreversibly cleaves sucrose into glucose and fructose.

As used herein, sucrose synthase is any enzyme that carries out the kinetically reversible transglycosylation of sucrose and UDP into fructose and UDPG.

As used herein, hexokinase is any enzyme that can phosphorylate either glucose or fructose.

As used herein, fructokinase is any enzyme that typically has a specific affinity for fructose.

As used herein, NDP-kinase is any enzyme that can maintain UDP levels for sucrose synthase reactions, even in the case of an ATP-specific fructokinase.

As used herein, glucose-6-phosphate 1-dehydrogenase is any enzyme that allows G6P resulting from hexose kinase activity to enter the pentose phosphate pathway.

As used herein, UDP-glucose dehydrogenase is any enzyme that allows UDPG formed by sucrose synthase to be utilized directly for cellulose or callose biosynthesis.

As used herein, phosphoglucomutase is any enzyme that is ubiquitous and reversible with commitments of G6P to either F6P or G1P resulting from fluxes in metabolites further along each pathway.

## **Agents**

### **(a) Nucleic Acid Molecules**

Agents of the present invention include plant nucleic acid molecules and more preferably include maize and soybean nucleic acid molecules and more preferably include nucleic acid

molecules of the maize genotypes B73 (Illinois Foundation Seeds, Champaign, Illinois U.S.A.), B73 x Mo17 (Illinois Foundation Seeds, Champaign, Illinois U.S.A.), DK604 (Dekalb Genetics, Dekalb, Illinois U.S.A.), H99 (Illinois Foundation Seeds, Champaign, Illinois U.S.A.), RX601 (Asgrow Seed Company, Des Moines, Iowa), Mo17 (Illinois Foundation Seeds, Champaign, Illinois U.S.A.), and soybean types Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa), C1944 (United States Department of Agriculture (USDA) Soybean Germplasm Collection, Urbana, Illinois U.S.A.), Cristalina (USDA Soybean Germplasm Collection, Urbana, Illinois U.S.A.), FT108 (Monsoy, Brazil), Hartwig (USDA Soybean Germplasm Collection, Urbana, Illinois U.S.A.), BW211S Null (Tohoku University, Morioka, Japan), PI507354 (USDA Soybean Germplasm Collection, Urbana, Illinois U.S.A.), Asgrow A4922 (Asgrow Seed Company, Des Moines, Iowa U.S.A.), PI227687 (USDA Soybean Germplasm Collection, Urbana, Illinois U.S.A.), PI229358 (USDA Soybean Germplasm Collection, Urbana, Illinois U.S.A.) and Asgrow A3237 (Asgrow Seed Company, Des Moines, Iowa U.S.A.).

A subset of the nucleic acid molecules of the present invention includes nucleic acid molecules that are marker molecules. Another subset of the nucleic acid molecules of the present invention include nucleic acid molecules that encode a protein or fragment thereof. Another subset of the nucleic acid molecules of the present invention are EST molecules.

Fragment nucleic acid molecules may encode significant portion(s) of, or indeed most of, these nucleic acid molecules. Alternatively, the fragments may comprise smaller oligonucleotides (having from about 15 to about 250 nucleotide residues and more preferably, about 15 to about 30 nucleotide residues).

As used herein, an agent, be it a naturally occurring molecule or otherwise may be “substantially purified,” if desired, such that one or more molecules that is or may be present in a



naturally occurring preparation containing that molecule will have been removed or will be present at a lower concentration than that at which it would normally be found.

The agents of the present invention will preferably be "biologically active" with respect to either a structural attribute, such as the capacity of a nucleic acid to hybridize to another nucleic acid molecule, or the ability of a protein to be bound by an antibody (or to compete with another molecule for such binding). Alternatively, such an attribute may be catalytic and thus involve the capacity of the agent to mediate a chemical reaction or response.

The agents of the present invention may also be recombinant. As used herein, the term recombinant means any agent (e.g. DNA, peptide etc.), that is, or results, however indirect, from human manipulation of a nucleic acid molecule.

It is understood that the agents of the present invention may be labeled with reagents that facilitate detection of the agent (e.g. fluorescent labels, Prober *et al.*, *Science* 238:336-340 (1987); Albarella *et al.*, EP 144914; chemical labels, Sheldon *et al.*, U.S. Patent 4,582,789; Albarella *et al.*, U.S. Patent 4,563,417; modified bases, Miyoshi *et al.*, EP 119448, all of which are hereby incorporated by reference in their entirety).

It is further understood, that the present invention provides recombinant bacterial, mammalian, microbial, insect, fungal and plant cells and viral constructs comprising the agents of the present invention (See, for example, Uses of the Agents of the Invention, Section (a) Plant Constructs and Plant Transformants; Section (b) Fungal Constructs and Fungal Transformants; Section (c) Mammalian Constructs and Transformed Mammalian Cells; Section (d) Insect Constructs and Transformed Insect Cells; and Section (e) Bacterial Constructs and Transformed Bacterial Cells).

Nucleic acid molecules or fragments thereof of the present invention are capable of specifically hybridizing to other nucleic acid molecules under certain circumstances. As used herein, two nucleic acid molecules are said to be capable of specifically hybridizing to one another if the two molecules are capable of forming an anti-parallel, double-stranded nucleic acid structure. A nucleic acid molecule is said to be the "complement" of another nucleic acid molecule if they exhibit complete complementarity. As used herein, molecules are said to exhibit "complete complementarity" when every nucleotide of one of the molecules is complementary to a nucleotide of the other. Two molecules are said to be "minimally complementary" if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under at least conventional "low-stringency" conditions. Similarly, the molecules are said to be "complementary" if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under conventional "high-stringency" conditions. Conventional stringency conditions are described by Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Press, Cold Spring Harbor, New York (1989) and by Haymes *et al.*, *Nucleic Acid Hybridization, A Practical Approach*, IRL Press, Washington, DC (1985), the entirety of which is herein incorporated by reference. Departures from complete complementarity are therefore permissible, as long as such departures do not completely preclude the capacity of the molecules to form a double-stranded structure. Thus, in order for a nucleic acid molecule to serve as a primer or probe it need only be sufficiently complementary in sequence to be able to form a stable double-stranded structure under the particular solvent and salt concentrations employed.

Appropriate stringency conditions which promote DNA hybridization, for example, 6.0 X sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 X SSC at 50°C,

are known to those skilled in the art or can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0 X SSC at 50°C to a high stringency of about 0.2 X SSC at 50°C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22°C, to high stringency conditions at about 65°C. Both temperature and salt may be varied, or either the temperature or the salt concentration may be held constant while the other variable is changed.

In a preferred embodiment, a nucleic acid of the present invention will specifically hybridize to one or more of the nucleic acid molecules set forth in SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof under moderately stringent conditions, for example at about 2.0 X SSC and about 65°C.

In a particularly preferred embodiment, a nucleic acid of the present invention will include those nucleic acid molecules that specifically hybridize to one or more of the nucleic acid molecules set forth in SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof under high stringency conditions such as 0.2 X SSC and about 65°C.

In one aspect of the present invention, the nucleic acid molecules of the present invention have one or more of the nucleic acid sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof. In another aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 90% sequence identity with one or more of the nucleic acid sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof. In a further aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 95% sequence identity with one or more of the nucleic acid sequences set forth in SEQ ID NO: 1 through SEQ ID NO:

2814 or complements thereof. In a more preferred aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 98% sequence identity with one or more of the nucleic acid sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof. In an even more preferred aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 99% sequence identity with one or more of the sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof.

In a further more preferred aspect of the present invention, one or more of the nucleic acid molecules of the present invention exhibit 100% sequence identity with a nucleic acid molecule present within MONN01, SATMON001 through SATMON031, SATMON033, SATMON034, SATMON~001, SATMONNN01, SATMONNN04 through SATMONNN006, CMz029 through CMz031, CMz033, CMz035 through CMz037, CMz039 through CMz042, CMz044 through CMz045, CMz047 through CMz050, SOYMON001 through SOYMON038, Soy51 through Soy56, Soy58 through Soy62, Soy65 through Soy66, Soy 68 through Soy73 and Soy76 through Soy77, Lib9, Lib22 through Lib25, Lib35, Lib80 through Lib81, Lib 144, Lib146, Lib147, Lib190, Lib3032 through Lib3036 and Lib3099 (Monsanto Company, St. Louis, Missouri U.S.A.).

**(i) Nucleic Acid Molecules Encoding Proteins or Fragments Thereof**

Nucleic acid molecules of the present invention can comprise sequences that encode a sucrose pathway protein or fragment thereof. Such proteins or fragments thereof include homologues of known proteins in other organisms.

In a preferred embodiment of the present invention, a maize or a soybean protein or fragment thereof of the present invention is a homologue of another plant protein. In another

preferred embodiment of the present invention, a maize or a soybean protein or fragment thereof of the present invention is a homologue of a fungal protein. In another preferred embodiment of the present invention, a maize or a soybean protein of the present invention is a homologue of a mammalian protein. In another preferred embodiment of the present invention, a maize or a soybean protein or fragment thereof of the present invention is a homologue of a bacterial protein. In another preferred embodiment of the present invention, a soybean protein or fragment thereof of the present invention is a homologue of a maize protein. In another preferred embodiment of the present invention, a maize protein homologue or fragment thereof of the present invention is a homologue of a soybean protein.

In a preferred embodiment of the present invention, the nucleic molecule of the present invention encodes a maize or a soybean protein or fragment thereof where a maize or a soybean protein exhibits a BLAST probability score of greater than  $1E-12$ , preferably a BLAST probability score of between about  $1E-30$  and about  $1E-12$ , even more preferably a BLAST probability score of greater than  $1E-30$  with its homologue.

In another preferred embodiment of the present invention, the nucleic acid molecule encoding a maize or a soybean protein or fragment thereof exhibits a % identity with its homologue of between about 25% and about 40%, more preferably of between about 40 and about 70%, even more preferably of between about 70% and about 90% and even more preferably between about 90% and 99%. In another preferred embodiment, of the present invention, a maize or a soybean protein or fragments thereof exhibits a % identity with its homologue of 100%.

In a preferred embodiment of the present invention, the nucleic molecule of the present invention encodes a maize or a soybean protein or fragment thereof where a maize or a soybean

protein exhibits a BLAST score of greater than 120, preferably a BLAST score of between about 1450 and about 120, even more preferably a BLAST score of greater than 1450 with its homologue.

Nucleic acid molecules of the present invention also include non-maize, non-soybean homologues. Preferred non-maize and soybean homologues are selected from the group consisting of alfalfa, *Arabidopsis*, barley, *Brassica*, broccoli, cabbage, citrus, cotton, garlic, oat, oilseed rape, onion, canola, flax, an ornamental plant, pea, peanut, pepper, potato, rice, rye, sorghum, strawberry, sugarcane, sugarbeet, tomato, wheat, poplar, pine, fir, eucalyptus, apple, lettuce, lentils, grape, banana, tea, turf grasses, sunflower, oil palm and *Phaseolus*.

In a preferred embodiment, nucleic acid molecules having SEQ ID NO: 1 through SEQ ID NO: 2814 or complements and fragments of either can be utilized to obtain such homologues.

The degeneracy of the genetic code, which allows different nucleic acid sequences to code for the same protein or peptide, is known in the literature. (U.S. Patent No. 4,757,006, the entirety of which is herein incorporated by reference).

In an aspect of the present invention, one or more of the nucleic acid molecules of the present invention differ in nucleic acid sequence from those encoding a maize or a soybean protein or fragment thereof in SEQ ID NO: 1 through SEQ ID NO: 2814 due to the degeneracy in the genetic code in that they encode the same protein but differ in nucleic acid sequence.

In another further aspect of the present invention, one or more of the nucleic acid molecules of the present invention differ in nucleic acid sequence from those encoding a maize or a soybean protein or fragment thereof in SEQ ID NO: 1 through SEQ ID NO: 2814 due to fact that the different nucleic acid sequence encodes a protein having one or more conservative amino acid residue. Examples of conservative substitutions are set forth in Table 1. It is

understood that codons capable of coding for such conservative substitutions are known in the art.

**Table 1**

<u>Original Residue</u>	<u>Conservative Substitutions</u>
Ala	Ser
Arg	Lys
Asn	Gln; His
Asp	Glu
Cys	Ser; Ala
Gln	Asn
Glu	Asp
Gly	Pro
His	Asn; Gln
Ile	Leu; Val
Leu	Ile; Val
Lys	Arg; Gln; Glu
Met	Leu; Ile
Phe	Met; Leu; Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp; Phe

Val

Ile; Leu

In a further aspect of the present invention, one or more of the nucleic acid molecules of the present invention differ in nucleic acid sequence from those encoding a maize or a soybean protein or fragment thereof set forth in SEQ ID NO: 1 through SEQ ID NO: 2814 or fragment thereof due to the fact that one or more codons encoding an amino acid has been substituted for a codon that encodes a nonessential substitution of the amino acid originally encoded.

Agents of the present invention include nucleic acid molecules that encode a maize or a soybean sucrose pathway protein or fragment thereof and particularly substantially purified nucleic acid molecules selected from the group consisting of a nucleic acid molecule that encodes a maize or a soybean triose phosphate isomerase protein or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate aldolase protein or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate protein or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 6-phosphate 2-kinase protein or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean phosphoglucosomerase protein or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase protein or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase protein or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean invertase protein or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean sucrose synthase protein or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean hexokinase protein or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructokinase protein or fragment



thereof, a nucleic acid molecule that encodes a maize or a soybean NDP-kinase protein or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean glucose-6-phosphate 1-dehydrogenase protein or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean phosphoglucomutase protein or fragment thereof and a nucleic acid molecule that encodes a maize or a soybean UDP-glucose pyrophosphorylase protein or fragment thereof.

Non-limiting examples of such nucleic acid molecules of the present invention are nucleic acid molecules comprising: SEQ ID NO: 1 through SEQ ID NO: 2814 or fragment thereof that encode for a sucrose pathway protein or fragment thereof, SEQ ID NO: 1 through SEQ ID NO: 206 and SEQ ID NO: 1538 through SEQ ID NO: 1707 or fragment thereof that encode for a triose phosphate isomerase protein or fragment thereof, SEQ ID NO: 207 through SEQ ID NO: 232 and SEQ ID NO: 1708 through SEQ ID NO: 2113 or fragment thereof that encode for a fructose 1,6-bisphosphate aldolase protein or fragment thereof, SEQ ID NO: 233 through SEQ ID NO: 258 and SEQ ID NO: 2114 through SEQ ID NO: 2162 or fragment thereof that encode for a fructose 1,6-bisphosphate protein or fragment thereof, SEQ ID NO: 259 through SEQ ID NO: 275 and SEQ ID NO: 2163 through SEQ ID NO: 2166 or fragment thereof that encode for a fructose 6-phosphate 2-kinase protein or fragment thereof, SEQ ID NO: 276 through SEQ ID NO: 340 and SEQ ID NO: 2167 through SEQ ID NO: 2182 or fragment thereof that encode for a phosphoglucosomerase protein or fragment thereof, SEQ ID NO: 341 through SEQ ID NO: 497 and SEQ ID NO: 2183 through SEQ ID NO: 2241 or fragment thereof that encode for a vacuolar H<sup>+</sup> translocating-pyrophosphatase protein or fragment thereof, SEQ ID NO: 498 through SEQ ID NO: 507 and SEQ ID NO: 2442 or fragment thereof that encode for a pyrophosphate-dependent fructose-6-phosphate phosphotransferase protein or fragment thereof, SEQ ID NO: 508 through SEQ ID NO: 510 and SEQ ID NO: 2243 through SEQ ID NO: 2254 or

fragment thereof that encode for an invertase protein or fragment thereof, SEQ ID NO: 511 through SEQ ID NO: 1086 and SEQ ID NO: 2255 through SEQ ID NO: 2590 or fragment thereof that encode for a sucrose synthase protein or fragment thereof, SEQ ID NO: 1087 through SEQ ID NO: 1135 and SEQ ID NO: 2591 through SEQ ID NO: 2634 or fragment thereof that encode for a hexokinase protein or fragment thereof, SEQ ID NO: 1136 through SEQ ID NO: 1215 and SEQ ID NO: 2635 through SEQ ID NO: 2678 or fragment thereof that encode for a fructokinase protein or fragment thereof, SEQ ID NO: 1216 through SEQ ID NO: 1251 and SEQ ID NO: 2679 through SEQ ID NO: 2681 or fragment thereof that encode for a NDP-kinase protein or fragment thereof, SEQ ID NO: 1252 through SEQ ID NO: 1254 and SEQ ID NO: 2682 through SEQ ID NO: 2689 or fragment thereof that encode for a glucose-6-phosphate 1-dehydrogenase protein or fragment thereof, SEQ ID NO: 1255 through SEQ ID NO: 1360 and SEQ ID NO: 2690 through SEQ ID NO: 2740 or fragment thereof that encode for a phosphoglucomutase protein or fragment thereof and SEQ ID NO: 1361 through SEQ ID NO: 1537 and SEQ ID NO: 2741 through SEQ ID NO: 2814 or fragment thereof that encode for an UDP-glucose pyrophosphorylase protein or fragment thereof.

A nucleic acid molecule of the present invention can also encode a homologue of a maize or a soybean triose phosphate isomerase or fragment thereof, a maize or a soybean fructose 1,6-bisphosphate aldolase or fragment thereof, a maize or a soybean fructose 1,6-bisphosphate or fragment thereof, a maize or a soybean fructose 6-phosphate 2-kinase or fragment thereof, a maize or a soybean phosphoglucoisomerase or fragment thereof, a maize or a soybean vacuolar  $H^+$  translocating-pyrophosphatase or fragment thereof, a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase or fragment thereof, a maize or a soybean invertase or fragment thereof, a maize or a soybean sucrose synthase or fragment thereof, a maize

or a soybean hexokinase or fragment thereof, a maize or a soybean fructokinase or fragment thereof, a maize or a soybean NDP-kinase or fragment thereof, a maize or a soybean glucose-6-phosphate 1-dehydrogenase or fragment thereof, a maize or a soybean phosphoglucomutase or fragment thereof and a maize or a soybean UDP-glucose pyrophosphorylase or fragment thereof. As used herein a homologue protein molecule or fragment thereof is a counterpart protein molecule or fragment thereof in a second species (*e.g.*, maize triose phosphate isomerase protein is a homologue of soybean triose phosphate isomerase protein).

**(ii) Nucleic Acid Molecule Markers and Probes**

One aspect of the present invention concerns markers that include nucleic acid molecules SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof or fragments of either that can act as markers or other nucleic acid molecules of the present invention that can act as markers.. Genetic markers of the present invention include “dominant” or “codominant” markers “Codominant markers” reveal the presence of two or more alleles (two per diploid individual) at a locus. “Dominant markers” reveal the presence of only a single allele per locus. The presence of the dominant marker phenotype (*e.g.*, a band of DNA) is an indication that one allele is present in either the homozygous or heterozygous condition. The absence of the dominant marker phenotype (*e.g.* absence of a DNA band) is merely evidence that “some other” undefined allele is present. In the case of populations where individuals are predominantly homozygous and loci are predominately dimorphic, dominant and codominant markers can be equally valuable. As populations become more heterozygous and multi-allelic, codominant markers often become more informative of the genotype than dominant markers. Marker molecules can be, for example, capable of detecting polymorphisms such as single nucleotide polymorphisms (SNPs).

SNPs are single base changes in genomic DNA sequence. They occur at greater frequency and are spaced with a greater uniformity throughout a genome than other reported forms of polymorphism. The greater frequency and uniformity of SNPs means that there is greater probability that such a polymorphism will be found near or in a genetic locus of interest than would be the case for other polymorphisms. SNPs are located in protein-coding regions and noncoding regions of a genome. Some of these SNPs may result in defective or variant protein expression (e.g., as a result of mutations or defective splicing). Analysis (genotyping) of characterized SNPs can require only a plus/minus assay rather than a lengthy measurement, permitting easier automation.

SNPs can be characterized using any of a variety of methods. Such methods include the direct or indirect sequencing of the site, the use of restriction enzymes (Botstein *et al.*, *Am. J. Hum. Genet.* 32:314-331 (1980), the entirety of which is herein incorporated reference; Konieczny and Ausubel, *Plant J.* 4:403-410 (1993), the entirety of which is herein incorporated by reference), enzymatic and chemical mismatch assays (Myers *et al.*, *Nature* 313:495-498 (1985), the entirety of which is herein incorporated by reference), allele-specific PCR (Newton *et al.*, *Nucl. Acids Res.* 17:2503-2516 (1989), the entirety of which is herein incorporated by reference; Wu *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 86:2757-2760 (1989), the entirety of which is herein incorporated by reference), ligase chain reaction (Barany, *Proc. Natl. Acad. Sci. (U.S.A.)* 88:189-193 (1991), the entirety of which is herein incorporated by reference), single-strand conformation polymorphism analysis (Labrune *et al.*, *Am. J. Hum. Genet.* 48: 1115-1120 (1991), the entirety of which is herein incorporated by reference), primer-directed nucleotide incorporation assays (Kuppuswami *et al.*, *Proc. Natl. Acad. Sci. USA* 88:1143-1147 (1991), the entirety of which is herein incorporated by reference), dideoxy fingerprinting (Sarkar *et al.*,

Genomics 13:441-443 (1992), the entirety of which is herein incorporated by reference), solid-phase ELISA-based oligonucleotide ligation assays (Nikiforov *et al.*, *Nucl. Acids Res.* 22:4167-4175 (1994), the entirety of which is herein incorporated by reference), oligonucleotide fluorescence-quenching assays (Livak *et al.*, *PCR Methods Appl.* 4:357-362 (1995), the entirety of which is herein incorporated by reference), 5'-nuclease allele-specific hybridization TaqMan assay (Livak *et al.*, *Nature Genet.* 9:341-342 (1995), the entirety of which is herein incorporated by reference), template-directed dye-terminator incorporation (TDI) assay (Chen and Kwok, *Nucl. Acids Res.* 25:347-353 (1997), the entirety of which is herein incorporated by reference), allele-specific molecular beacon assay (Tyagi *et al.*, *Nature Biotech.* 16: 49-53 (1998), the entirety of which is herein incorporated by reference), PinPoint assay (Haff and Smirnov, *Genome Res.* 7: 378-388 (1997), the entirety of which is herein incorporated by reference) and dCAPS analysis (Neff *et al.*, *Plant J.* 14:387-392 (1998), the entirety of which is herein incorporated by reference).

Additional markers, such as AFLP markers, RFLP markers and RAPD markers, can be utilized (Walton, *Seed World* 22-29 (July, 1993), the entirety of which is herein incorporated by reference; Burow and Blake, *Molecular Dissection of Complex Traits*, 13-29, Paterson (ed.), CRC Press, New York (1988), the entirety of which is herein incorporated by reference). DNA markers can be developed from nucleic acid molecules using restriction endonucleases, the PCR and/or DNA sequence information. RFLP markers result from single base changes or insertions/deletions. These codominant markers are highly abundant in plant genomes, have a medium level of polymorphism and are developed by a combination of restriction endonuclease digestion and Southern blotting hybridization. CAPS are similarly developed from restriction nuclease digestion but only of specific PCR products. These markers are also codominant, have

a medium level of polymorphism and are highly abundant in the genome. The CAPS result from single base changes and insertions/deletions.

Another marker type, RAPDs, are developed from DNA amplification with random primers and result from single base changes and insertions/deletions in plant genomes. They are dominant markers with a medium level of polymorphisms and are highly abundant. AFLP markers require using the PCR on a subset of restriction fragments from extended adapter primers. These markers are both dominant and codominant are highly abundant in genomes and exhibit a medium level of polymorphism.

SSRs require DNA sequence information. These codominant markers result from repeat length changes, are highly polymorphic and do not exhibit as high a degree of abundance in the genome as CAPS, AFLPs and RAPDs SNPs also require DNA sequence information. These codominant markers result from single base substitutions. They are highly abundant and exhibit a medium of polymorphism (Rafalski *et al.*, In: *Nonmammalian Genomic Analysis*, Birren and Lai (ed.), Academic Press, San Diego, CA, pp. 75-134 (1996), the entirety of which is herein incorporated by reference). It is understood that a nucleic acid molecule of the present invention may be used as a marker.

A PCR probe is a nucleic acid molecule capable of initiating a polymerase activity while in a double-stranded structure to with another nucleic acid. Various methods for determining the structure of PCR probes and PCR techniques exist in the art. Computer generated searches using programs such as Primer3 ([www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi](http://www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi)), STSPipeline ([www-genome.wi.mit.edu/cgi-bin/www-STSPipeline](http://www-genome.wi.mit.edu/cgi-bin/www-STSPipeline)), or GeneUp (Pesole *et al.*, *BioTechniques* 25:112-123 (1998) the entirety of which is herein incorporated by reference), for example, can be used to identify potential PCR primers.

It is understood that a fragment of one or more of the nucleic acid molecules of the present invention may be a probe and specifically a PCR probe.

**(b) Protein and Peptide Molecules**

A class of agents comprises one or more of the protein or fragments thereof or peptide molecules encoded by SEQ ID NO: 1 through SEQ ID NO: 2814 or one or more of the protein or fragment thereof and peptide molecules encoded by other nucleic acid agents of the present invention. As used herein, the term "protein molecule" or "peptide molecule" includes any molecule that comprises five or more amino acids. It is well known in the art that proteins may undergo modification, including post-translational modifications, such as, but not limited to, disulfide bond formation, glycosylation, phosphorylation, or oligomerization. Thus, as used herein, the term "protein molecule" or "peptide molecule" includes any protein molecule that is modified by any biological or non-biological process. The terms "amino acid" and "amino acids" refer to all naturally occurring L-amino acids. This definition is meant to include norleucine, ornithine, homocysteine and homoserine.

Non-limiting examples of the protein or fragment thereof of the present invention include a maize or a soybean sucrose pathway protein or fragment thereof; a maize or a soybean triose phosphate isomerase or fragment thereof, a maize or a soybean fructose 1,6-bisphosphate aldolase or fragment thereof, a maize or a soybean fructose 1,6-bisphosphate or fragment thereof, a maize or a soybean fructose 6-phosphate 2-kinase or fragment thereof, a maize or a soybean phosphoglucisomerase or fragment thereof, a maize or a soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase or fragment thereof, a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase or fragment thereof, a maize or a soybean invertase or fragment thereof, a maize or a soybean sucrose synthase or fragment thereof, a maize or a soybean

hexokinase or fragment thereof, a maize or a soybean fructokinase or fragment thereof, a maize or a soybean NDP-kinase or fragment thereof, a maize or a soybean glucose-6-phosphate 1-dehydrogenase or fragment thereof, a maize or a soybean phosphoglucomutase or fragment thereof and a maize or a soybean UDP-glucose pyrophosphorylase or fragment thereof.

Non-limiting examples of the protein or fragment molecules of the present invention are a sucrose pathway protein or fragment thereof encoded by: SEQ ID NO: 1 through SEQ ID NO: 2814 or fragment thereof that encode for a sucrose pathway protein or fragment thereof, SEQ ID NO: 1 through SEQ ID NO: 206 and SEQ ID NO: 1538 through SEQ ID NO: 1707 or fragment thereof that encode for a triose phosphate isomerase protein or fragment thereof, SEQ ID NO: 207 through SEQ ID NO: 232 and SEQ ID NO: 1708 through SEQ ID NO: 2113 or fragment thereof that encode for a fructose 1,6-bisphosphate aldolase protein or fragment thereof, SEQ ID NO: 233 through SEQ ID NO: 258 and SEQ ID NO: 2114 through SEQ ID NO: 2162 or fragment thereof that encode for a fructose 1,6-bisphosphate protein or fragment thereof, SEQ ID NO: 259 through SEQ ID NO: 275 and SEQ ID NO: 2163 through SEQ ID NO: 2166 or fragment thereof that encode for a fructose 6-phosphate 2-kinase protein or fragment thereof, SEQ ID NO: 276 through SEQ ID NO: 340 and SEQ ID NO: 2167 through SEQ ID NO: 2182 or fragment thereof that encode for a phosphoglucosomerase protein or fragment thereof, SEQ ID NO: 341 through SEQ ID NO: 497 and SEQ ID NO: 2183 through SEQ ID NO: 2241 or fragment thereof that encode for a vacuolar H<sup>+</sup> translocating-pyrophosphatase protein or fragment thereof, SEQ ID NO: 498 through SEQ ID NO: 507 and SEQ ID NO: 2442 or fragment thereof that encode for a pyrophosphate-dependent fructose-6-phosphate phosphotransferase protein or fragment thereof, SEQ ID NO: 508 through SEQ ID NO: 510 and SEQ ID NO: 2243 through SEQ ID NO: 2254 or fragment thereof that encode for an invertase protein or fragment thereof,



SEQ ID NO: 511 through SEQ ID NO: 1086 and SEQ ID NO: 2255 through SEQ ID NO: 2590 or fragment thereof that encode for a sucrose synthase protein or fragment thereof, SEQ ID NO: 1087 through SEQ ID NO: 1135 and SEQ ID NO: 2591 through SEQ ID NO: 2634 or fragment thereof that encode for a hexokinase protein or fragment thereof, SEQ ID NO: 1136 through SEQ ID NO: 1215 and SEQ ID NO: 2635 through SEQ ID NO: 2678 or fragment thereof that encode for a fructokinase protein or fragment thereof, SEQ ID NO: 1216 through SEQ ID NO: 1251 and SEQ ID NO: 2679 through SEQ ID NO: 2681 or fragment thereof that encode for a NDP-kinase protein or fragment thereof, SEQ ID NO: 1252 through SEQ ID NO: 1254 and SEQ ID NO: 2682 through SEQ ID NO: 2689 or fragment thereof that encode for a glucose-6-phosphate 1-dehydrogenase protein or fragment thereof, SEQ ID NO: 1255 through SEQ ID NO: 1360 and SEQ ID NO: 2690 through SEQ ID NO: 2740 or fragment thereof that encode for a phosphoglucomutase protein or fragment thereof and SEQ ID NO: 1361 through SEQ ID NO: 1537 and SEQ ID NO: 2741 through SEQ ID NO: 2814 or fragment thereof that encode for an UDP-glucose pyrophosphorylase protein or fragment thereof.

One or more of the protein or fragment of peptide molecules may be produced via chemical synthesis, or more preferably, by expressing in a suitable bacterial or eucaryotic host. Suitable methods for expression are described by Sambrook *et al.*, (In: *Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, New York* (1989)), or similar texts. For example, the protein may be expressed in, for example, Uses of the Agents of the Invention, Section (a) Plant Constructs and Plant Transformants; Section (b) Fungal Constructs and Fungal Transformants; Section (c) Mammalian Constructs and Transformed Mammalian Cells; Section (d) Insect Constructs and Transformed Insect Cells; and Section (e) Bacterial Constructs and Transformed Bacterial Cells.

A "protein fragment" is a peptide or polypeptide molecule whose amino acid sequence comprises a subset of the amino acid sequence of that protein. A protein or fragment thereof that comprises one or more additional peptide regions not derived from that protein is a "fusion" protein. Such molecules may be derivatized to contain carbohydrate or other moieties (such as keyhole limpet hemocyanin, etc.). Fusion protein or peptide molecules of the present invention are preferably produced via recombinant means.

Another class of agents comprise protein or peptide molecules or fragments or fusions thereof encoded by SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof in which conservative, non-essential or non-relevant amino acid residues have been added, replaced or deleted. Computerized means for designing modifications in protein structure are known in the art (Dahiyat and Mayo, *Science* 278:82-87 (1997), the entirety of which is herein incorporated by reference).

The protein molecules of the present invention include plant homologue proteins. An example of such a homologue is a homologue protein of a non-maize or non-soybean plant species, that include but not limited to alfalfa, *Arabidopsis*, barley, *Brassica*, broccoli, cabbage, citrus, cotton, garlic, oat, oilseed rape, onion, canola, flax, an ornamental plant, pea, peanut, pepper, potato, rice, rye, sorghum, strawberry, sugarcane, sugarbeet, tomato, wheat, poplar, pine, fir, eucalyptus, apple, lettuce, lentils, grape, banana, tea, turf grasses, sunflower, oil palm, *Phaseolus* etc. Particularly preferred non-maize or non-soybean for use for the isolation of homologs would include, *Arabidopsis*, barley, cotton, oat, oilseed rape, rice, canola, ornamentals, sugarcane, sugarbeet, tomato, potato, wheat and turf grasses. Such a homologue can be obtained by any of a variety of methods. Most preferably, as indicated above, one or more of the disclosed sequences (SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof) will be

used to define a pair of primers that may be used to isolate the homologue-encoding nucleic acid molecules from any desired species. Such molecules can be expressed to yield homologues by recombinant means.

### **(c) Antibodies**

One aspect of the present invention concerns antibodies, single-chain antigen binding molecules, or other proteins that specifically bind to one or more of the protein or peptide molecules of the present invention and their homologues, fusions or fragments. Such antibodies may be used to quantitatively or qualitatively detect the protein or peptide molecules of the present invention. As used herein, an antibody or peptide is said to “specifically bind” to a protein or peptide molecule of the present invention if such binding is not competitively inhibited by the presence of non-related molecules.

Nucleic acid molecules that encode all or part of the protein of the present invention can be expressed, via recombinant means, to yield protein or peptides that can in turn be used to elicit antibodies that are capable of binding the expressed protein or peptide. Such antibodies may be used in immunoassays for that protein. Such protein-encoding molecules, or their fragments may be a “fusion” molecule (i.e., a part of a larger nucleic acid molecule) such that, upon expression, a fusion protein is produced. It is understood that any of the nucleic acid molecules of the present invention may be expressed, via recombinant means, to yield proteins or peptides encoded by these nucleic acid molecules.

The antibodies that specifically bind proteins and protein fragments of the present invention may be polyclonal or monoclonal and may comprise intact immunoglobulins, or antigen binding portions of immunoglobulins fragments (such as (F(ab'), F(ab')<sub>2</sub>), or single-chain

immunoglobulins producible, for example, via recombinant means. It is understood that practitioners are familiar with the standard resource materials which describe specific conditions and procedures for the construction, manipulation and isolation of antibodies (*see*, for example, Harlow and Lane, In: *Antibodies: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, New York (1988), the entirety of which is herein incorporated by reference).

Murine monoclonal antibodies are particularly preferred. BALB/c mice are preferred for this purpose, however, equivalent strains may also be used. The animals are preferably immunized with approximately 25 µg of purified protein (or fragment thereof) that has been emulsified in a suitable adjuvant (such as TiterMax adjuvant (Vaxcel, Norcross, GA)). Immunization is preferably conducted at two intramuscular sites, one intraperitoneal site and one subcutaneous site at the base of the tail. An additional i.v. injection of approximately 25 µg of antigen is preferably given in normal saline three weeks later. After approximately 11 days following the second injection, the mice may be bled and the blood screened for the presence of anti-protein or peptide antibodies. Preferably, a direct binding Enzyme-Linked Immunoassay (ELISA) is employed for this purpose.

More preferably, the mouse having the highest antibody titer is given a third i.v. injection of approximately 25 µg of the same protein or fragment. The splenic leukocytes from this animal may be recovered 3 days later and then permitted to fuse, most preferably, using polyethylene glycol, with cells of a suitable myeloma cell line (such as, for example, the P3X63Ag8.653 myeloma cell line). Hybridoma cells are selected by culturing the cells under "HAT" (hypoxanthine-aminopterin-thymine) selection for about one week. The resulting clones may then be screened for their capacity to produce monoclonal antibodies ("mAbs"), preferably by direct ELISA.

In one embodiment, anti-protein or peptide monoclonal antibodies are isolated using a fusion of a protein or peptide of the present invention, or conjugate of a protein or peptide of the present invention, as immunogens. Thus, for example, a group of mice can be immunized using a fusion protein emulsified in Freund's complete adjuvant (*e.g.* approximately 50 µg of antigen per immunization). At three week intervals, an identical amount of antigen is emulsified in Freund's incomplete adjuvant and used to immunize the animals. Ten days following the third immunization, serum samples are taken and evaluated for the presence of antibody. If antibody titers are too low, a fourth booster can be employed. Polysera capable of binding the protein or peptide can also be obtained using this method.

In a preferred procedure for obtaining monoclonal antibodies, the spleens of the above-described immunized mice are removed, disrupted and immune splenocytes are isolated over a ficoll gradient. The isolated splenocytes are fused, using polyethylene glycol with BALB/c-derived HGPRT (hypoxanthine guanine phosphoribosyl transferase) deficient P3x63xAg8.653 plasmacytoma cells. The fused cells are plated into 96 well microtiter plates and screened for hybridoma fusion cells by their capacity to grow in culture medium supplemented with hypoxanthine, aminopterin and thymidine for approximately 2-3 weeks.

Hybridoma cells that arise from such incubation are preferably screened for their capacity to produce an immunoglobulin that binds to a protein of interest. An indirect ELISA may be used for this purpose. In brief, the supernatants of hybridomas are incubated in microtiter wells that contain immobilized protein. After washing, the titer of bound immunoglobulin can be determined using, for example, a goat anti-mouse antibody conjugated to horseradish peroxidase. After additional washing, the amount of immobilized enzyme is determined (for example through the use of a chromogenic substrate). Such screening is performed as quickly as possible

after the identification of the hybridoma in order to ensure that a desired clone is not overgrown by non-secreting neighbor cells. Desirably, the fusion plates are screened several times since the rates of hybridoma growth vary. In a preferred sub-embodiment, a different antigenic form may be used to screen the hybridoma. Thus, for example, the splenocytes may be immunized with one immunogen, but the resulting hybridomas can be screened using a different immunogen. It is understood that any of the protein or peptide molecules of the present invention may be used to raise antibodies.

As discussed below, such antibody molecules or their fragments may be used for diagnostic purposes. Where the antibodies are intended for diagnostic purposes, it may be desirable to derivatize them, for example with a ligand group (such as biotin) or a detectable marker group (such as a fluorescent group, a radioisotope or an enzyme).

The ability to produce antibodies that bind the protein or peptide molecules of the present invention permits the identification of mimetic compounds of those molecules. A “mimetic compound” is a compound that is not that compound, or a fragment of that compound, but which nonetheless exhibits an ability to specifically bind to antibodies directed against that compound.

It is understood that any of the agents of the present invention can be substantially purified and/or be biologically active and/or recombinant.

### **Uses of the Agents of the Invention**

Nucleic acid molecules and fragments thereof of the present invention may be employed to obtain other nucleic acid molecules from the same species (e.g., ESTs or fragment thereof from maize may be utilized to obtain other nucleic acid molecules from maize). Such nucleic acid molecules include the nucleic acid molecules that encode the complete coding sequence of a protein and promoters and flanking sequences of such molecules. In addition, such nucleic acid

6636T01-03ZEE60

molecules include nucleic acid molecules that encode for other isozymes or gene family members. Such molecules can be readily obtained by using the above-described nucleic acid molecules or fragments thereof to screen cDNA or genomic libraries obtained from maize or soybean. Methods for forming such libraries are well known in the art.

Nucleic acid molecules and fragments thereof of the present invention may also be employed to obtain nucleic acid homologues. Such homologues include the nucleic acid molecule of other plants or other organisms (*e.g.*, alfalfa, *Arabidopsis*, barley, *Brassica*, broccoli, cabbage, citrus, cotton, garlic, oat, oilseed rape, onion, canola, flax, an ornamental plant, pea, peanut, pepper, potato, rice, rye, sorghum, strawberry, sugarcane, sugarbeet, tomato, wheat, poplar, pine, fir, eucalyptus, apple, lettuce, lentils, grape, banana, tea, turf grasses, sunflower, oil palm, *Phaseolus*, etc.) including the nucleic acid molecules that encode, in whole or in part, protein homologues of other plant species or other organisms, sequences of genetic elements such as promoters and transcriptional regulatory elements. Such molecules can be readily obtained by using the above-described nucleic acid molecules or fragments thereof to screen cDNA or genomic libraries obtained from such plant species. Methods for forming such libraries are well known in the art. Such homologue molecules may differ in their nucleotide sequences from those found in one or more of SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof because complete complementarity is not needed for stable hybridization. The nucleic acid molecules of the present invention therefore also include molecules that, although capable of specifically hybridizing with the nucleic acid molecules, may lack "complete complementarity."

Any of a variety of methods may be used to obtain one or more of the above-described nucleic acid molecules (*Zamechik et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 83:4143-4146 (1986), the entirety of which is herein incorporated by reference; *Goodchild et al.*, *Proc. Natl. Acad. Sci.*

(U.S.A.) 85:5507-5511 (1988), the entirety of which is herein incorporated by reference; Wickstrom *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 85:1028-1032 (1988), the entirety of which is herein incorporated by reference; Holt *et al.*, *Molec. Cell. Biol.* 8:963-973 (1988), the entirety of which is herein incorporated by reference; Gerwitz *et al.*, *Science* 242:1303-1306 (1988), the entirety of which is herein incorporated by reference; Anfossi *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 86:3379-3383 (1989), the entirety of which is herein incorporated by reference; Becker *et al.*, *EMBO J.* 8:3685-3691 (1989); the entirety of which is herein incorporated by reference). Automated nucleic acid synthesizers may be employed for this purpose. In lieu of such synthesis, the disclosed nucleic acid molecules may be used to define a pair of primers that can be used with the polymerase chain reaction (Mullis *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 51:263-273 (1986); Erlich *et al.*, European Patent 50,424; European Patent 84,796; European Patent 258,017; European Patent 237,362; Mullis, European Patent 201,184; Mullis *et al.*, U.S. Patent 4,683,202; Erlich, U.S. Patent 4,582,788; and Saiki *et al.*, U.S. Patent 4,683,194, all of which are herein incorporated by reference in their entirety) to amplify and obtain any desired nucleic acid molecule or fragment.

Promoter sequence(s) and other genetic elements, including but not limited to transcriptional regulatory flanking sequences, associated with one or more of the disclosed nucleic acid sequences can also be obtained using the disclosed nucleic acid sequence provided herein. In one embodiment, such sequences are obtained by incubating EST nucleic acid molecules or preferably fragments thereof with members of genomic libraries (*e.g.* maize and soybean) and recovering clones that hybridize to the EST nucleic acid molecule or fragment thereof. In a second embodiment, methods of "chromosome walking," or inverse PCR may be used to obtain such sequences (Frohman *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 85:8998-9002



(1988); Ohara *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 86:5673-5677 (1989); Pang *et al.*, *Biotechniques* 22:1046-1048 (1977); Huang *et al.*, *Methods Mol. Biol.* 69:89-96 (1997); Huang *et al.*, *Method Mol. Biol.* 67:287-294 (1997); Benkel *et al.*, *Genet. Anal.* 13:123-127 (1996); Hartl *et al.*, *Methods Mol. Biol.* 58:293-301 (1996), all of which are herein incorporated by reference in their entirety).

The nucleic acid molecules of the present invention may be used to isolate promoters of cell enhanced, cell specific, tissue enhanced, tissue specific, developmentally or environmentally regulated expression profiles. Isolation and functional analysis of the 5' flanking promoter sequences of these genes from genomic libraries, for example, using genomic screening methods and PCR techniques would result in the isolation of useful promoters and transcriptional regulatory elements. These methods are known to those of skill in the art and have been described (See, for example, Birren *et al.*, *Genome Analysis: Analyzing DNA*, 1, (1997), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., the entirety of which is herein incorporated by reference). Promoters obtained utilizing the nucleic acid molecules of the present invention could also be modified to affect their control characteristics. Examples of such modifications would include but are not limited to enhanced sequences as reported in Uses of the Agents of the Invention, Section (a) Plant Constructs and Plant Transformants. Such genetic elements could be used to enhance gene expression of new and existing traits for crop improvements.

In one sub-aspect, such an analysis is conducted by determining the presence and/or identity of polymorphism(s) by one or more of the nucleic acid molecules of the present invention and more preferably one or more of the EST nucleic acid molecule or fragment thereof which are associated with a phenotype, or a predisposition to that phenotype.

Any of a variety of molecules can be used to identify such polymorphism(s). In one embodiment, one or more of the EST nucleic acid molecules (or a sub-fragment thereof) may be employed as a marker nucleic acid molecule to identify such polymorphism(s). Alternatively, such polymorphisms can be detected through the use of a marker nucleic acid molecule or a marker protein that is genetically linked to (i.e., a polynucleotide that co-segregates with) such polymorphism(s).

In an alternative embodiment, such polymorphisms can be detected through the use of a marker nucleic acid molecule that is physically linked to such polymorphism(s). For this purpose, marker nucleic acid molecules comprising a nucleotide sequence of a polynucleotide located within 1mb of the polymorphism(s) and more preferably within 100kb of the polymorphism(s) and most preferably within 10kb of the polymorphism(s) can be employed.

The genomes of animals and plants naturally undergo spontaneous mutation in the course of their continuing evolution (Gusella, *Ann. Rev. Biochem.* 55:831-854 (1986)). A “polymorphism” is a variation or difference in the sequence of the gene or its flanking regions that arises in some of the members of a species. The variant sequence and the “original” sequence co-exist in the species’ population. In some instances, such co-existence is in stable or quasi-stable equilibrium.

A polymorphism is thus said to be “allelic,” in that, due to the existence of the polymorphism, some members of a species may have the original sequence (i.e., the original “allele”) whereas other members may have the variant sequence (i.e., the variant “allele”). In the simplest case, only one variant sequence may exist and the polymorphism is thus said to be di-allelic. In other cases, the species’ population may contain multiple alleles and the polymorphism is termed tri-allelic, etc. A single gene may have multiple different unrelated

polymorphisms. For example, it may have a di-allelic polymorphism at one site and a multi-allelic polymorphism at another site.

The variation that defines the polymorphism may range from a single nucleotide variation to the insertion or deletion of extended regions within a gene. In some cases, the DNA sequence variations are in regions of the genome that are characterized by short tandem repeats (STRs) that include tandem di- or tri-nucleotide repeated motifs of nucleotides. Polymorphisms characterized by such tandem repeats are referred to as "variable number tandem repeat" ("VNTR") polymorphisms. VNTRs have been used in identity analysis (Weber, U.S. Patent 5,075,217; Armour *et al.*, *FEBS Lett.* 307:113-115 (1992); Jones *et al.*, *Eur. J. Haematol.* 39:144-147 (1987); Horn *et al.*, PCT Patent Application WO91/14003; Jeffreys, European Patent Application 370,719; Jeffreys, U.S. Patent 5,175,082; Jeffreys *et al.*, *Amer. J. Hum. Genet.* 39:11-24 (1986); Jeffreys *et al.*, *Nature* 316:76-79 (1985); Gray *et al.*, *Proc. R. Acad. Soc. Lond.* 243:241-253 (1991); Moore *et al.*, *Genomics* 10:654-660 (1991); Jeffreys *et al.*, *Anim. Genet.* 18:1-15 (1987); Hillel *et al.*, *Anim. Genet.* 20:145-155 (1989); Hillel *et al.*, *Genet.* 124:783-789 (1990), all of which are herein incorporated by reference in their entirety).

The detection of polymorphic sites in a sample of DNA may be facilitated through the use of nucleic acid amplification methods. Such methods specifically increase the concentration of polynucleotides that span the polymorphic site, or include that site and sequences located either distal or proximal to it. Such amplified molecules can be readily detected by gel electrophoresis or other means.

The most preferred method of achieving such amplification employs the polymerase chain reaction ("PCR") (Mullis *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 51:263-273 (1986); Erlich *et al.*, European Patent Appln. 50,424; European Patent Appln. 84,796; European

Patent Application 258,017; European Patent Appln. 237,362; Mullis, European Patent Appln. 201,184; Mullis *et al.*, U.S. Patent No. 4,683,202; Erlich, U.S. Patent No. 4,582,788; and Saiki *et al.*, U.S. Patent No. 4,683,194), using primer pairs that are capable of hybridizing to the proximal sequences that define a polymorphism in its double-stranded form.

In lieu of PCR, alternative methods, such as the "Ligase Chain Reaction" ("LCR") may be used (Barany, *Proc. Natl. Acad. Sci. (U.S.A.)* 88:189-193 (1991), the entirety of which is herein incorporated by reference). LCR uses two pairs of oligonucleotide probes to exponentially amplify a specific target. The sequences of each pair of oligonucleotides is selected to permit the pair to hybridize to abutting sequences of the same strand of the target. Such hybridization forms a substrate for a template-dependent ligase. As with PCR, the resulting products thus serve as a template in subsequent cycles and an exponential amplification of the desired sequence is obtained.

LCR can be performed with oligonucleotides having the proximal and distal sequences of the same strand of a polymorphic site. In one embodiment, either oligonucleotide will be designed to include the actual polymorphic site of the polymorphism. In such an embodiment, the reaction conditions are selected such that the oligonucleotides can be ligated together only if the target molecule either contains or lacks the specific nucleotide that is complementary to the polymorphic site present on the oligonucleotide. Alternatively, the oligonucleotides may be selected such that they do not include the polymorphic site (see, Segev, PCT Application WO 90/01069, the entirety of which is herein incorporated by reference).

The "Oligonucleotide Ligation Assay" ("OLA") may alternatively be employed (Landegren *et al.*, *Science* 241:1077-1080 (1988), the entirety of which is herein incorporated by reference). The OLA protocol uses two oligonucleotides which are designed to be capable of

hybridizing to abutting sequences of a single strand of a target. OLA, like LCR, is particularly suited for the detection of point mutations. Unlike LCR, however, OLA results in "linear" rather than exponential amplification of the target sequence.

Nickerson *et al.*, have described a nucleic acid detection assay that combines attributes of PCR and OLA (Nickerson *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 87:8923-8927 (1990), the entirety of which is herein incorporated by reference). In this method, PCR is used to achieve the exponential amplification of target DNA, which is then detected using OLA. In addition to requiring multiple and separate, processing steps, one problem associated with such combinations is that they inherit all of the problems associated with PCR and OLA.

Schemes based on ligation of two (or more) oligonucleotides in the presence of nucleic acid having the sequence of the resulting "di-oligonucleotide", thereby amplifying the di-oligonucleotide, are also known (Wu *et al.*, *Genomics* 4:560-569 (1989), the entirety of which is herein incorporated by reference) and may be readily adapted to the purposes of the present invention.

Other known nucleic acid amplification procedures, such as allele-specific oligomers, branched DNA technology, transcription-based amplification systems, or isothermal amplification methods may also be used to amplify and analyze such polymorphisms (Malek *et al.*, U.S. Patent 5,130,238; Davey *et al.*, European Patent Application 329,822; Schuster *et al.*, U.S. Patent 5,169,766; Miller *et al.*, PCT Patent Application WO 89/06700; Kwoh *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 86:1173-1177 (1989); Gingeras *et al.*, PCT Patent Application WO 88/10315; Walker *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 89:392-396 (1992), all of which are herein incorporated by reference in their entirety).

The identification of a polymorphism can be determined in a variety of ways. By correlating the presence or absence of it in a plant with the presence or absence of a phenotype, it is possible to predict the phenotype of that plant. If a polymorphism creates or destroys a restriction endonuclease cleavage site, or if it results in the loss or insertion of DNA (e.g., a VNTR polymorphism), it will alter the size or profile of the DNA fragments that are generated by digestion with that restriction endonuclease. As such, individuals that possess a variant sequence can be distinguished from those having the original sequence by restriction fragment analysis. Polymorphisms that can be identified in this manner are termed "restriction fragment length polymorphisms" ("RFLPs"). RFLPs have been widely used in human and plant genetic analyses (Glassberg, UK Patent Application 2135774; Skolnick *et al.*, *Cytogen. Cell Genet.* 32:58-67 (1982); Botstein *et al.*, *Ann. J. Hum. Genet.* 32:314-331 (1980); Fischer *et al.*, (PCT Application WO90/13668); Uhlen, PCT Application WO90/11369).

Polymorphisms can also be identified by Single Strand Conformation Polymorphism (SSCP) analysis. SSCP is a method capable of identifying most sequence variations in a single strand of DNA, typically between 150 and 250 nucleotides in length (Elles, *Methods in Molecular Medicine: Molecular Diagnosis of Genetic Diseases*, Humana Press (1996), the entirety of which is herein incorporated by reference); Orita *et al.*, *Genomics* 5:874-879 (1989), the entirety of which is herein incorporated by reference). Under denaturing conditions a single strand of DNA will adopt a conformation that is uniquely dependent on its sequence conformation. This conformation usually will be different, even if only a single base is changed. Most conformations have been reported to alter the physical configuration or size sufficiently to be detectable by electrophoresis. A number of protocols have been described for SSCP including, but not limited to, Lee *et al.*, *Anal. Biochem.* 205:289-293 (1992), the entirety of

which is herein incorporated by reference; Suzuki *et al.*, *Anal. Biochem.* 192:82-84 (1991), the entirety of which is herein incorporated by reference; Lo *et al.*, *Nucleic Acids Research* 20:1005-1009 (1992), the entirety of which is herein incorporated by reference; Sarkar *et al.*, *Genomics* 13:441-443 (1992), the entirety of which is herein incorporated by reference. It is understood that one or more of the nucleic acids of the present invention, may be utilized as markers or probes to detect polymorphisms by SSCP analysis.

Polymorphisms may also be found using a DNA fingerprinting technique called amplified fragment length polymorphism (AFLP), which is based on the selective PCR amplification of restriction fragments from a total digest of genomic DNA to profile that DNA (Vos *et al.*, *Nucleic Acids Res.* 23:4407-4414 (1995), the entirety of which is herein incorporated by reference). This method allows for the specific co-amplification of high numbers of restriction fragments, which can be visualized by PCR without knowledge of the nucleic acid sequence.

AFLP employs basically three steps. Initially, a sample of genomic DNA is cut with restriction enzymes and oligonucleotide adapters are ligated to the restriction fragments of the DNA. The restriction fragments are then amplified using PCR by using the adapter and restriction sequence as target sites for primer annealing. The selective amplification is achieved by the use of primers that extend into the restriction fragments, amplifying only those fragments in which the primer extensions match the nucleotide flanking the restriction sites. These amplified fragments are then visualized on a denaturing polyacrylamide gel.

AFLP analysis has been performed on *Salix* (Beismann *et al.*, *Mol. Ecol.* 6:989-993 (1997), the entirety of which is herein incorporated by reference), *Acinetobacter* (Janssen *et al.*, *Int. J. Syst. Bacteriol.* 47:1179-1187 (1997), the entirety of which is herein incorporated by

reference), *Aeromonas popoffi* (Huys *et al.*, *Int. J. Syst. Bacteriol.* 47:1165-1171 (1997), the entirety of which is herein incorporated by reference), rice (McCouch *et al.*, *Plant Mol. Biol.* 35:89-99 (1997), the entirety of which is herein incorporated by reference; Nandi *et al.*, *Mol. Gen. Genet.* 255:1-8 (1997), the entirety of which is herein incorporated by reference; Cho *et al.*, *Genome* 39:373-378 (1996), the entirety of which is herein incorporated by reference), barley (*Hordeum vulgare*)(Simons *et al.*, *Genomics* 44:61-70 (1997), the entirety of which is herein incorporated by reference; Waugh *et al.*, *Mol. Gen. Genet.* 255:311-321 (1997), the entirety of which is herein incorporated by reference; Qi *et al.*, *Mol. Gen. Genet.* 254:330-336 (1997), the entirety of which is herein incorporated by reference; Becker *et al.*, *Mol. Gen. Genet.* 249:65-73 (1995), the entirety of which is herein incorporated by reference), potato (Van der Voort *et al.*, *Mol. Gen. Genet.* 255:438-447 (1997), the entirety of which is herein incorporated by reference; Meksem *et al.*, *Mol. Gen. Genet.* 249:74-81 (1995), the entirety of which is herein incorporated by reference), *Phytophthora infestans* (Van der Lee *et al.*, *Fungal Genet. Biol.* 21:278-291 (1997), the entirety of which is herein incorporated by reference), *Bacillus anthracis* (Keim *et al.*, *J. Bacteriol.* 179:818-824 (1997), the entirety of which is herein incorporated by reference), *Astragalus cremnophylax* (Travis *et al.*, *Mol. Ecol.* 5:735-745 (1996), the entirety of which is herein incorporated by reference), *Arabidopsis* (Cnops *et al.*, *Mol. Gen. Genet.* 253:32-41 (1996), the entirety of which is herein incorporated by reference), *Escherichia coli* (Lin *et al.*, *Nucleic Acids Res.* 24:3649-3650 (1996), the entirety of which is herein incorporated by reference), *Aeromonas* (Huys *et al.*, *Int. J. Syst. Bacteriol.* 46:572-580 (1996), the entirety of which is herein incorporated by reference), nematode (Folkertsma *et al.*, *Mol. Plant Microbe Interact.* 9:47-54 (1996), the entirety of which is herein incorporated by reference), tomato (Thomas *et al.*, *Plant J.* 8:785-794 (1995), the entirety of which is herein incorporated by reference) and human (Latorra



*et al.*, *PCR Methods Appl.* 3:351-358 (1994), the entirety of which is herein incorporated by reference). AFLP analysis has also been used for fingerprinting mRNA (Money *et al.*, *Nucleic Acids Res.* 24:2616-2617 (1996), the entirety of which is herein incorporated by reference; Bachem *et al.*, *Plant J.* 9:745-753 (1996), the entirety of which is herein incorporated by reference). It is understood that one or more of the nucleic acids of the present invention, may be utilized as markers or probes to detect polymorphisms by AFLP analysis or for fingerprinting RNA.

Polymorphisms may also be found using random amplified polymorphic DNA (RAPD) (Williams *et al.*, *Nucl. Acids Res.* 18:6531-6535 (1990), the entirety of which is herein incorporated by reference) and cleaveable amplified polymorphic sequences (CAPS) (Lyamichev *et al.*, *Science* 260:778-783 (1993), the entirety of which is herein incorporated by reference). It is understood that one or more of the nucleic acid molecules of the present invention, may be utilized as markers or probes to detect polymorphisms by RAPD or CAPS analysis.

Through genetic mapping, a fine scale linkage map can be developed using DNA markers and, then, a genomic DNA library of large-sized fragments can be screened with molecular markers linked to the desired trait. Molecular markers are advantageous for agronomic traits that are otherwise difficult to tag, such as resistance to pathogens, insects and nematodes, tolerance to abiotic stress, quality parameters and quantitative traits such as high yield potential.

The essential requirements for marker-assisted selection in a plant breeding program are: (1) the marker(s) should co-segregate or be closely linked with the desired trait; (2) an efficient means of screening large populations for the molecular marker(s) should be available; and (3) the screening technique should have high reproducibility across laboratories and preferably be economical to use and be user-friendly.

The genetic linkage of marker molecules can be established by a gene mapping model such as, without limitation, the flanking marker model reported by Lander and Botstein, *Genetics* 121:185-199 (1989) and the interval mapping, based on maximum likelihood methods described by Lander and Botstein, *Genetics* 121:185-199 (1989) and implemented in the software package MAPMAKER/QTL (Lincoln and Lander, *Mapping Genes Controlling Quantitative Traits Using MAPMAKER/QTL*, Whitehead Institute for Biomedical Research, Massachusetts, (1990). Additional software includes Qgene, Version 2.23 (1996), Department of Plant Breeding and Biometry, 266 Emerson Hall, Cornell University, Ithaca, NY, the manual of which is herein incorporated by reference in its entirety). Use of Qgene software is a particularly preferred approach.

A maximum likelihood estimate (MLE) for the presence of a marker is calculated, together with an MLE assuming no QTL effect, to avoid false positives. A  $\log_{10}$  of an odds ratio (LOD) is then calculated as:  $LOD = \log_{10}(MLE \text{ for the presence of a QTL} / MLE \text{ given no linked QTL})$ .

The LOD score essentially indicates how much more likely the data are to have arisen assuming the presence of a QTL than in its absence. The LOD threshold value for avoiding a false positive with a given confidence, say 95%, depends on the number of markers and the length of the genome. Graphs indicating LOD thresholds are set forth in Lander and Botstein, *Genetics* 121:185-199 (1989) the entirety of which is herein incorporated by reference and further described by Arús and Moreno-González, *Plant Breeding*, Hayward *et al.*, (eds.) Chapman & Hall, London, pp. 314-331 (1993), the entirety of which is herein incorporated by reference.

Additional models can be used. Many modifications and alternative approaches to interval mapping have been reported, including the use non-parametric methods (Kruglyak and Lander, *Genetics* 139:1421-1428 (1995), the entirety of which is herein incorporated by reference). Multiple regression methods or models can be also be used, in which the trait is regressed on a large number of markers (Jansen, *Biometrics in Plant Breeding*, van Oijen and Jansen (eds.), Proceedings of the Ninth Meeting of the Eucarpia Section Biometrics in Plant Breeding, The Netherlands, pp. 116-124 (1994); Weber and Wricke, *Advances in Plant Breeding*, Blackwell, Berlin, 16 (1994), both of which is herein incorporated by reference in their entirety). Procedures combining interval mapping with regression analysis, whereby the phenotype is regressed onto a single putative QTL at a given marker interval and at the same time onto a number of markers that serve as 'cofactors,' have been reported by Jansen and Stam, *Genetics* 136:1447-1455 (1994), the entirety of which is herein incorporated by reference and Zeng, *Genetics* 136:1457-1468 (1994) the entirety of which is herein incorporated by reference. Generally, the use of cofactors reduces the bias and sampling error of the estimated QTL positions (Utz and Melchinger, *Biometrics in Plant Breeding*, van Oijen and Jansen (eds.) Proceedings of the Ninth Meeting of the Eucarpia Section Biometrics in Plant Breeding, The Netherlands, pp.195-204 (1994), the entirety of which is herein incorporated by reference, thereby improving the precision and efficiency of QTL mapping (Zeng, *Genetics* 136:1457-1468 (1994)). These models can be extended to multi-environment experiments to analyze genotype-environment interactions (Jansen *et al.*, *Theo. Appl. Genet.* 91:33-37 (1995), the entirety of which is herein incorporated by reference).

Selection of an appropriate mapping populations is important to map construction. The choice of appropriate mapping population depends on the type of marker systems employed

(Tanksley *et al.*, *Molecular mapping plant chromosomes. Chromosome structure and function: Impact of new concepts*, Gustafson and Appels (eds.), Plenum Press, New York, pp 157-173 (1988), the entirety of which is herein incorporated by reference). Consideration must be given to the source of parents (adapted vs. exotic) used in the mapping population. Chromosome pairing and recombination rates can be severely disturbed (suppressed) in wide crosses (adapted x exotic) and generally yield greatly reduced linkage distances. Wide crosses will usually provide segregating populations with a relatively large array of polymorphisms when compared to progeny in a narrow cross (adapted x adapted).

An  $F_2$  population is the first generation of selfing after the hybrid seed is produced. Usually a single  $F_1$  plant is selfed to generate a population segregating for all the genes in Mendelian (1:2:1) fashion. Maximum genetic information is obtained from a completely classified  $F_2$  population using a codominant marker system (Mather, *Measurement of Linkage in Heredity*, Methuen and Co., (1938), the entirety of which is herein incorporated by reference). In the case of dominant markers, progeny tests (e.g.  $F_3$ ,  $BCF_2$ ) are required to identify the heterozygotes, thus making it equivalent to a completely classified  $F_2$  population. However, this procedure is often prohibitive because of the cost and time involved in progeny testing. Progeny testing of  $F_2$  individuals is often used in map construction where phenotypes do not consistently reflect genotype (e.g. disease resistance) or where trait expression is controlled by a QTL. Segregation data from progeny test populations (e.g.  $F_3$  or  $BCF_2$ ) can be used in map construction. Marker-assisted selection can then be applied to cross progeny based on marker-trait map associations ( $F_2$ ,  $F_3$ ), where linkage groups have not been completely disassociated by recombination events (i.e., maximum disequilibrium).

Recombinant inbred lines (RIL) (genetically related lines; usually  $>F_3$ , developed from continuously selfing  $F_2$  lines towards homozygosity) can be used as a mapping population. Information obtained from dominant markers can be maximized by using RIL because all loci are homozygous or nearly so. Under conditions of tight linkage (i.e., about  $<10\%$  recombination), dominant and co-dominant markers evaluated in RIL populations provide more information per individual than either marker type in backcross populations (Reiter *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 89:1477-1481 (1992), the entirety of which is herein incorporated by reference). However, as the distance between markers becomes larger (i.e., loci become more independent), the information in RIL populations decreases dramatically when compared to codominant markers.

Backcross populations (e.g., generated from a cross between a successful variety (recurrent parent) and another variety (donor parent) carrying a trait not present in the former) can be utilized as a mapping population. A series of backcrosses to the recurrent parent can be made to recover most of its desirable traits. Thus a population is created consisting of individuals nearly like the recurrent parent but each individual carries varying amounts or mosaic of genomic regions from the donor parent. Backcross populations can be useful for mapping dominant markers if all loci in the recurrent parent are homozygous and the donor and recurrent parent have contrasting polymorphic marker alleles (Reiter *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 89:1477-1481 (1992)). Information obtained from backcross populations using either codominant or dominant markers is less than that obtained from  $F_2$  populations because one, rather than two, recombinant gametes are sampled per plant. Backcross populations, however, are more informative (at low marker saturation) when compared to RILs as the distance between linked loci increases in RIL populations (i.e. about  $15\%$  recombination). Increased

recombination can be beneficial for resolution of tight linkages, but may be undesirable in the construction of maps with low marker saturation.

Near-isogenic lines (NIL) created by many backcrosses to produce an array of individuals that are nearly identical in genetic composition except for the trait or genomic region under interrogation can be used as a mapping population. In mapping with NILs, only a portion of the polymorphic loci are expected to map to a selected region.

Bulk segregant analysis (BSA) is a method developed for the rapid identification of linkage between markers and traits of interest (Michelmore *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 88:9828-9832 (1991), the entirety of which is herein incorporated by reference). In BSA, two bulked DNA samples are drawn from a segregating population originating from a single cross. These bulks contain individuals that are identical for a particular trait (resistant or susceptible to particular disease) or genomic region but arbitrary at unlinked regions (i.e. heterozygous). Regions unlinked to the target region will not differ between the bulked samples of many individuals in BSA.

It is understood that one or more of the nucleic acid molecules of the present invention may be used as molecular markers. It is also understood that one or more of the protein molecules of the present invention may be used as molecular markers.

In accordance with this aspect of the present invention, a sample nucleic acid is obtained from plants cells or tissues. Any source of nucleic acid may be used. Preferably, the nucleic acid is genomic DNA. The nucleic acid is subjected to restriction endonuclease digestion. For example, one or more nucleic acid molecule or fragment thereof of the present invention can be used as a probe in accordance with the above-described polymorphic methods. The polymorphism obtained in this approach can then be cloned to identify the mutation at the coding

region which alters the protein's structure or regulatory region of the gene which affects its expression level.

In an aspect of the present invention, one or more of the nucleic molecules of the present invention are used to determine the level (i.e., the concentration of mRNA in a sample, etc.) in a plant (preferably maize or soybean) or pattern (i.e., the kinetics of expression, rate of decomposition, stability profile, etc.) of the expression of a protein encoded in part or whole by one or more of the nucleic acid molecule of the present invention (collectively, the "Expression Response" of a cell or tissue). As used herein, the Expression Response manifested by a cell or tissue is said to be "altered" if it differs from the Expression Response of cells or tissues of plants not exhibiting the phenotype. To determine whether a Expression Response is altered, the Expression Response manifested by the cell or tissue of the plant exhibiting the phenotype is compared with that of a similar cell or tissue sample of a plant not exhibiting the phenotype. As will be appreciated, it is not necessary to re-determine the Expression Response of the cell or tissue sample of plants not exhibiting the phenotype each time such a comparison is made; rather, the Expression Response of a particular plant may be compared with previously obtained values of normal plants. As used herein, the phenotype of the organism is any of one or more characteristics of an organism (e.g. disease resistance, pest tolerance, environmental tolerance such as tolerance to abiotic stress, male sterility, quality improvement or yield etc.). A change in genotype or phenotype may be transient or permanent. Also as used herein, a tissue sample is any sample that comprises more than one cell. In a preferred aspect, a tissue sample comprises cells that share a common characteristic (e.g. derived from root, seed, flower, leaf, stem or pollen etc.).

In one aspect of the present invention, an evaluation can be conducted to determine whether a particular mRNA molecule is present. One or more of the nucleic acid molecules of the present invention, preferably one or more of the EST nucleic acid molecules or fragments thereof of the present invention are utilized to detect the presence or quantity of the mRNA species. Such molecules are then incubated with cell or tissue extracts of a plant under conditions sufficient to permit nucleic acid hybridization. The detection of double-stranded probe-mRNA hybrid molecules is indicative of the presence of the mRNA; the amount of such hybrid formed is proportional to the amount of mRNA. Thus, such probes may be used to ascertain the level and extent of the mRNA production in a plant's cells or tissues. Such nucleic acid hybridization may be conducted under quantitative conditions (thereby providing a numerical value of the amount of the mRNA present). Alternatively, the assay may be conducted as a qualitative assay that indicates either that the mRNA is present, or that its level exceeds a user set, predefined value.

A principle of *in situ* hybridization is that a labeled, single-stranded nucleic acid probe will hybridize to a complementary strand of cellular DNA or RNA and, under the appropriate conditions, these molecules will form a stable hybrid. When nucleic acid hybridization is combined with histological techniques, specific DNA or RNA sequences can be identified within a single cell. An advantage of *in situ* hybridization over more conventional techniques for the detection of nucleic acids is that it allows an investigator to determine the precise spatial population (Angerer *et al.*, *Dev. Biol.* 101:477-484 (1984), the entirety of which is herein incorporated by reference; Angerer *et al.*, *Dev. Biol.* 112:157-166 (1985), the entirety of which is herein incorporated by reference; Dixon *et al.*, *EMBO J.* 10:1317-1324 (1991), the entirety of which is herein incorporated by reference). *In situ* hybridization may be used to measure the



steady-state level of RNA accumulation. It is a sensitive technique and RNA sequences present in as few as 5-10 copies per cell can be detected (Hardin *et al.*, *J. Mol. Biol.* 202:417-431 (1989), the entirety of which is herein incorporated by reference). A number of protocols have been devised for *in situ* hybridization, each with tissue preparation, hybridization and washing conditions (Meyerowitz, *Plant Mol. Biol. Rep.* 5:242-250 (1987), the entirety of which is herein incorporated by reference; Cox and Goldberg, In: *Plant Molecular Biology: A Practical Approach*, Shaw (ed.), pp 1-35, IRL Press, Oxford (1988), the entirety of which is herein incorporated by reference; Raikhel *et al.*, *In situ RNA hybridization in plant tissues*, In: *Plant Molecular Biology Manual*, vol. B9:1-32, Kluwer Academic Publisher, Dordrecht, Belgium (1989), the entirety of which is herein incorporated by reference).

*In situ* hybridization also allows for the localization of proteins within a tissue or cell (Wilkinson, *In Situ Hybridization*, Oxford University Press, Oxford (1992), the entirety of which is herein incorporated by reference; Langdale, *In Situ Hybridization* In: *The Maize Handbook*, Freeling and Walbot (eds.), pp 165-179, Springer-Verlag, New York (1994), the entirety of which is herein incorporated by reference). It is understood that one or more of the molecules of the present invention, preferably one or more of the EST nucleic acid molecules or fragments thereof of the present invention or one or more of the antibodies of the present invention may be utilized to detect the level or pattern of a sucrose pathway protein or mRNA thereof by *in situ* hybridization.

Fluorescent *in situ* hybridization allows the localization of a particular DNA sequence along a chromosome which is useful, among other uses, for gene mapping, following chromosomes in hybrid lines or detecting chromosomes with translocations, transversions or deletions. *In situ* hybridization has been used to identify chromosomes in several plant species

(Griffor *et al.*, *Plant Mol. Biol.* 17:101-109 (1991), the entirety of which is herein incorporated by reference; Gustafson *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 87:1899-1902 (1990), herein incorporated by reference; Mukai and Gill, *Genome* 34:448-452 (1991), the entirety of which is herein incorporated by reference; Schwarzacher and Heslop-Harrison, *Genome* 34:317-323 (1991); Wang *et al.*, *Jpn. J. Genet.* 66:313-316 (1991), the entirety of which is herein incorporated by reference; Parra and Windle, *Nature Genetics* 5:17-21 (1993), the entirety of which is herein incorporated by reference). It is understood that the nucleic acid molecules of the present invention may be used as probes or markers to localize sequences along a chromosome.

Another method to localize the expression of a molecule is tissue printing. Tissue printing provides a way to screen, at the same time on the same membrane many tissue sections from different plants or different developmental stages. Tissue-printing procedures utilize films designed to immobilize proteins and nucleic acids. In essence, a freshly cut section of a tissue is pressed gently onto nitrocellulose paper, nylon membrane or polyvinylidene difluoride membrane. Such membranes are commercially available (*e.g.* Millipore, Bedford, Massachusetts U.S.A.). The contents of the cut cell transfer onto the membrane and the contents and are immobilized to the membrane. The immobilized contents form a latent print that can be visualized with appropriate probes. When a plant tissue print is made on nitrocellulose paper, the cell walls leave a physical print that makes the anatomy visible without further treatment (Varner and Taylor, *Plant Physiol.* 91:31-33 (1989), the entirety of which is herein incorporated by reference).

Tissue printing on substrate films is described by Daoust, *Exp. Cell Res.* 12:203-211 (1957), the entirety of which is herein incorporated by reference, who detected amylase, protease, ribonuclease and deoxyribonuclease in animal tissues using starch, gelatin and agar films. These

techniques can be applied to plant tissues (Yomo and Taylor, *Planta* 112:35-43 (1973); the entirety of which is herein incorporated by reference; Harris and Chrispeels, *Plant Physiol.* 56:292-299 (1975), the entirety of which is herein incorporated by reference). Advances in membrane technology have increased the range of applications of Daoust's tissue-printing techniques allowing (Cassab and Varner, *J. Cell. Biol.* 105:2581-2588 (1987), the entirety of which is herein incorporated by reference) the histochemical localization of various plant enzymes and deoxyribonuclease on nitrocellulose paper and nylon (Spruce *et al.*, *Phytochemistry* 26:2901-2903 (1987), the entirety of which is herein incorporated by reference; Barres *et al.*, *Neuron* 5:527-544 (1990), the entirety of which is herein incorporated by reference; Reid and Pont-Lezica, *Tissue Printing: Tools for the Study of Anatomy, Histochemistry and Gene Expression*, Academic Press, New York, New York (1992), the entirety of which is herein incorporated by reference; Reid *et al.*, *Plant Physiol.* 93:160-165 (1990), the entirety of which is herein incorporated by reference; Ye *et al.*, *Plant J.* 1:175-183 (1991), the entirety of which is herein incorporated by reference).

It is understood that one or more of the molecules of the present invention, preferably one or more of the EST nucleic acid molecules or fragments thereof of the present invention or one or more of the antibodies of the present invention may be utilized to detect the presence or quantity of a sucrose pathway protein by tissue printing.

Further it is also understood that any of the nucleic acid molecules of the present invention may be used as marker nucleic acids and or probes in connection with methods that require probes or marker nucleic acids. As used herein, a probe is an agent that is utilized to determine an attribute or feature (e.g. presence or absence, location, correlation, etc.) of a molecule, cell, tissue or plant. As used herein, a marker nucleic acid is a nucleic acid molecule

that is utilized to determine an attribute or feature (*e.g.*, presence or absence, location, correlation, etc.) or a molecule, cell, tissue or plant.

A microarray-based method for high-throughput monitoring of plant gene expression may be utilized to measure gene-specific hybridization targets. This 'chip'-based approach involves using microarrays of nucleic acid molecules as gene-specific hybridization targets to quantitatively measure expression of the corresponding plant genes (Schena *et al.*, *Science* 270:467-470 (1995), the entirety of which is herein incorporated by reference; Shalon, Ph.D. Thesis, Stanford University (1996), the entirety of which is herein incorporated by reference). Every nucleotide in a large sequence can be queried at the same time. Hybridization can be used to efficiently analyze nucleotide sequences.

Several microarray methods have been described. One method compares the sequences to be analyzed by hybridization to a set of oligonucleotides representing all possible subsequences (Bains and Smith, *J. Theor. Biol.* 135:303-307 (1989), the entirety of which is herein incorporated by reference). A second method hybridizes the sample to an array of oligonucleotide or cDNA molecules. An array consisting of oligonucleotides complementary to subsequences of a target sequence can be used to determine the identity of a target sequence, measure its amount and detect differences between the target and a reference sequence. Nucleic acid molecules microarrays may also be screened with protein molecules or fragments thereof to determine nucleic acid molecules that specifically bind protein molecules or fragments thereof.

The microarray approach may be used with polypeptide targets (U.S. Patent No. 5,445,934; U.S. Patent No: 5,143,854; U.S. Patent No. 5,079,600; U.S. Patent No. 4,923,901, all of which are herein incorporated by reference in their entirety). Essentially, polypeptides are synthesized on a substrate (microarray) and these polypeptides can be screened with either

protein molecules or fragments thereof or nucleic acid molecules in order to screen for either protein molecules or fragments thereof or nucleic acid molecules that specifically bind the target polypeptides. (Fodor *et al.*, *Science* 251:767-773 (1991), the entirety of which is herein incorporated by reference). It is understood that one or more of the nucleic acid molecules or protein or fragments thereof of the present invention may be utilized in a microarray based method.

In a preferred embodiment of the present invention microarrays may be prepared that comprise nucleic acid molecules where such nucleic acid molecules encode at least one, preferably at least two, more preferably at least three or preferably at least four, preferably at least five, preferably at least six, preferably at least seven, preferably at least eight, preferably at least nine, preferably at least ten, preferably at least eleven, preferably at least twelve, preferably at least thirteen, preferably at least fourteen preferably at least fifteen sucrose pathway enzymes. In a preferred embodiment the nucleic acid molecules are selected from the group consisting of a nucleic acid molecule that encodes a maize or a soybean triose phosphate isomerase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate aldolase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructose 6-phosphate 2-kinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean phosphoglucisomerase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean invertase enzyme

or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean sucrose synthase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean hexokinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean fructokinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean NDP-kinase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean glucose-6-phosphate 1-dehydrogenase enzyme or fragment thereof, a nucleic acid molecule that encodes a maize or a soybean phosphoglucomutase enzyme or fragment thereof and a nucleic acid molecule that encodes a maize or a soybean UDP-glucose pyrophosphorylase enzyme or fragment thereof.

Site directed mutagenesis may be utilized to modify nucleic acid sequences, particularly as it is a technique that allows one or more of the amino acids encoded by a nucleic acid molecule to be altered (e.g. a threonine to be replaced by a methionine). Three basic methods for site directed mutagenesis are often employed. These are cassette mutagenesis (Wells *et al.*, *Gene* 34:315-323 (1985), the entirety of which is herein incorporated by reference), primer extension (Gilliam *et al.*, *Gene* 12:129-137 (1980), the entirety of which is herein incorporated by reference; Zoller and Smith, *Methods Enzymol.* 100:468-500 (1983), the entirety of which is herein incorporated by reference; Dalbadie-McFarland *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 79:6409-6413 (1982), the entirety of which is herein incorporated by reference) and methods based upon PCR (Scharf *et al.*, *Science* 233:1076-1078 (1986), the entirety of which is herein incorporated by reference; Higuchi *et al.*, *Nucleic Acids Res.* 16:7351-7367 (1988), the entirety of which is herein incorporated by reference). Site directed mutagenesis approaches are also described in European Patent 0 385 962, the entirety of which is herein incorporated by reference; European Patent 0 359 472, the entirety of which is herein incorporated by reference;

and PCT Patent Application WO 93/07278, the entirety of which is herein incorporated by reference.

Site directed mutagenesis strategies have been applied to plants for both *in vitro* as well as *in vivo* site directed mutagenesis (Lanz *et al.*, *J. Biol. Chem.* 266:9971-9976 (1991), the entirety of which is herein incorporated by reference; Kovgan and Zhdanov, *Biotekhnologiya* 5:148-154; No. 207160n, Chemical Abstracts 110:225 (1989), the entirety of which is herein incorporated by reference; Ge *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 86:4037-4041 (1989), the entirety of which is herein incorporated by reference; Zhu *et al.*, *J. Biol. Chem.* 271:18494-18498 (1996), the entirety of which is herein incorporated by reference; Chu *et al.*, *Biochemistry* 33:6150-6157 (1994), the entirety of which is herein incorporated by reference; Small *et al.*, *EMBO J.* 11:1291-1296 (1992), the entirety of which is herein incorporated by reference; Cho *et al.*, *Mol. Biotechnol.* 8:13-16 (1997), the entirety of which is herein incorporated by reference; Kita *et al.*, *J. Biol. Chem.* 271:26529-26535 (1996), the entirety of which is herein incorporated by reference; Jin *et al.*, *Mol. Microbiol.* 7:555-562 (1993), the entirety of which is herein incorporated by reference; Hatfield and Vierstra, *J. Biol. Chem.* 267:14799-14803 (1992), the entirety of which is herein incorporated by reference; Zhao *et al.*, *Biochemistry* 31:5093-5099 (1992), the entirety of which is herein incorporated by reference).

Any of the nucleic acid molecules of the present invention may either be modified by site directed mutagenesis or used as, for example, nucleic acid molecules that are used to target other nucleic acid molecules for modification. It is understood that mutants with more than one altered nucleotide can be constructed using techniques that practitioners are familiar with such as isolating restriction fragments and ligating such fragments into an expression vector (*see*, for

example, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (1989)).

Sequence-specific DNA-binding proteins play a role in the regulation of transcription. The isolation of recombinant cDNAs encoding these proteins facilitates the biochemical analysis of their structural and functional properties. Genes encoding such DNA-binding proteins have been isolated using classical genetics (Vollbrecht *et al.*, *Nature* 350: 241-243 (1991), the entirety of which is herein incorporated by reference) and molecular biochemical approaches, including the screening of recombinant cDNA libraries with antibodies (Landschulz *et al.*, *Genes Dev.* 2:786-800 (1988), the entirety of which is herein incorporated by reference) or DNA probes (Bodner *et al.*, *Cell* 55:505-518 (1988), the entirety of which is herein incorporated by reference). In addition, an *in situ* screening procedure has been used and has facilitated the isolation of sequence-specific DNA-binding proteins from various plant species (Gilmartin *et al.*, *Plant Cell* 4:839-849 (1992), the entirety of which is herein incorporated by reference; Schindler *et al.*, *EMBO J.* 11:1261-1273 (1992), the entirety of which is herein incorporated by reference). An *in situ* screening protocol does not require the purification of the protein of interest (Vinson *et al.*, *Genes Dev.* 2:801-806 (1988), the entirety of which is herein incorporated by reference; Singh *et al.*, *Cell* 52:415-423 (1988), the entirety of which is herein incorporated by reference).

Two steps may be employed to characterize DNA-protein interactions. The first is to identify promoter fragments that interact with DNA-binding proteins, to titrate binding activity, to determine the specificity of binding and to determine whether a given DNA-binding activity can interact with related DNA sequences (Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2<sup>nd</sup> edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York



(1989)). Electrophoretic mobility-shift assay is a widely used assay. The assay provides a rapid and sensitive method for detecting DNA-binding proteins based on the observation that the mobility of a DNA fragment through a nondenaturing, low-ionic strength polyacrylamide gel is retarded upon association with a DNA-binding protein (Fried and Crother, *Nucleic Acids Res.* 9:6505-6525 (1981), the entirety of which is herein incorporated by reference). When one or more specific binding activities have been identified, the exact sequence of the DNA bound by the protein may be determined. Several procedures for characterizing protein/DNA-binding sites are used, including methylation and ethylation interference assays (Maxam and Gilbert, *Methods Enzymol.* 65:499-560 (1980), the entirety of which is herein incorporated by reference; Wissman and Hillen, *Methods Enzymol.* 208:365-379 (1991), the entirety of which is herein incorporated by reference), footprinting techniques employing DNase I (Galas and Schmitz, *Nucleic Acids Res.* 5:3157-3170 (1978), the entirety of which is herein incorporated by reference), 1,10-phenanthroline-copper ion methods (Sigman *et al.*, *Methods Enzymol.* 208:414-433 (1991), the entirety of which is herein incorporated by reference) and hydroxyl radicals methods (Dixon *et al.*, *Methods Enzymol.* 208:414-433 (1991), the entirety of which is herein incorporated by reference). It is understood that one or more of the nucleic acid molecules of the present invention may be utilized to identify a protein or fragment thereof that specifically binds to a nucleic acid molecule of the present invention. It is also understood that one or more of the protein molecules or fragments thereof of the present invention may be utilized to identify a nucleic acid molecule that specifically binds to it.

A two-hybrid system is based on the fact that many cellular functions are carried out by proteins, such as transcription factors, that interact (physically) with one another. Two-hybrid systems have been used to probe the function of new proteins (Chien *et al.*, *Proc. Natl. Acad. Sci.*

(U.S.A.) 88:9578-9582 (1991) the entirety of which is herein incorporated by reference; Durfee *et al.*, *Genes Dev.* 7:555-569 (1993) the entirety of which is herein incorporated by reference; Choi *et al.*, *Cell* 78:499-512 (1994), the entirety of which is herein incorporated by reference; Kranz *et al.*, *Genes Dev.* 8:313-327 (1994), the entirety of which is herein incorporated by reference).

Interaction mating techniques have facilitated a number of two-hybrid studies of protein-protein interaction. Interaction mating has been used to examine interactions between small sets of tens of proteins (Finley and Brent, *Proc. Natl. Acad. Sci. (U.S.A.)* 91:12098-12984 (1994), the entirety of which is herein incorporated by reference), larger sets of hundreds of proteins (Bendixen *et al.*, *Nucl. Acids Res.* 22:1778-1779 (1994), the entirety of which is herein incorporated by reference) and to comprehensively map proteins encoded by a small genome (Bartel *et al.*, *Nature Genetics* 12:72-77 (1996), the entirety of which is herein incorporated by reference). This technique utilizes proteins fused to the DNA-binding domain and proteins fused to the activation domain. They are expressed in two different haploid yeast strains of opposite mating type and the strains are mated to determine if the two proteins interact. Mating occurs when haploid yeast strains come into contact and result in the fusion of the two haploids into a diploid yeast strain. An interaction can be determined by the activation of a two-hybrid reporter gene in the diploid strain. An advantage of this technique is that it reduces the number of yeast transformations needed to test individual interactions. It is understood that the protein-protein interactions of protein or fragments thereof of the present invention may be investigated using the two-hybrid system and that any of the nucleic acid molecules of the present invention that encode such proteins or fragments thereof may be used to transform yeast in the two-hybrid system.

**(a) Plant Constructs and Plant Transformants**



generally within the skill of the art (See, *Plant Molecular Biology: A Laboratory Manual*, Clark (ed.), Springer, New York (1997), the entirety of which is herein incorporated by reference).

A construct or vector may include a plant promoter to express the protein or protein fragment of choice. A number of promoters which are active in plant cells have been described in the literature. These include the nopaline synthase (NOS) promoter (Ebert *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 84:5745-5749 (1987), the entirety of which is herein incorporated by reference), the octopine synthase (OCS) promoter (which are carried on tumor-inducing plasmids of *Agrobacterium tumefaciens*), the caulimovirus promoters such as the cauliflower mosaic virus (CaMV) 19S promoter (Lawton *et al.*, *Plant Mol. Biol.* 9:315-324 (1987), the entirety of which is herein incorporated by reference) and the CAMV 35S promoter (Odell *et al.*, *Nature* 313:810-812 (1985), the entirety of which is herein incorporated by reference), the figwort mosaic virus 35S-promoter, the light-inducible promoter from the small subunit of ribulose-1,5-bis-phosphate carboxylase (ssRUBISCO), the Adh promoter (Walker *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 84:6624-6628 (1987), the entirety of which is herein incorporated by reference), the sucrose synthase promoter (Yang *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 87:4144-4148 (1990), the entirety of which is herein incorporated by reference), the R gene complex promoter (Chandler *et al.*, *The Plant Cell* 1:1175-1183 (1989), the entirety of which is herein incorporated by reference) and the chlorophyll a/b binding protein gene promoter, etc. These promoters have been used to create DNA constructs which have been expressed in plants; see, e.g., PCT publication WO 84/02913, herein incorporated by reference in its entirety.

Promoters which are known or are found to cause transcription of DNA in plant cells can be used in the present invention. Such promoters may be obtained from a variety of sources such as plants and plant viruses. It is preferred that the particular promoter selected should be capable

of causing sufficient expression to result in the production of an effective amount of the sucrose pathway protein to cause the desired phenotype. In addition to promoters that are known to cause transcription of DNA in plant cells, other promoters may be identified for use in the current invention by screening a plant cDNA library for genes which are selectively or preferably expressed in the target tissues or cells.

For the purpose of expression in source tissues of the plant, such as the leaf, seed, root or stem, it is preferred that the promoters utilized in the present invention have relatively high expression in these specific tissues. For this purpose, one may choose from a number of promoters for genes with tissue- or cell-specific or -enhanced expression. Examples of such promoters reported in the literature include the chloroplast glutamine synthetase GS2 promoter from pea (Edwards *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 87:3459-3463 (1990), herein incorporated by reference in its entirety), the chloroplast fructose-1,6-bisphosphatase (FBPase) promoter from wheat (Lloyd *et al.*, *Mol. Gen. Genet.* 225:209-216 (1991), herein incorporated by reference in its entirety), the nuclear photosynthetic ST-LS1 promoter from potato (Stockhaus *et al.*, *EMBO J.* 8:2445-2451 (1989), herein incorporated by reference in its entirety), the serine/threonine kinase (PAL) promoter and the glucoamylase (CHS) promoter from *Arabidopsis thaliana*. Also reported to be active in photosynthetically active tissues are the ribulose-1,5-bisphosphate carboxylase (RbcS) promoter from eastern larch (*Larix laricina*), the promoter for the *cab* gene, *cab6*, from pine (Yamamoto *et al.*, *Plant Cell Physiol.* 35:773-778 (1994), herein incorporated by reference in its entirety), the promoter for the *Cab-1* gene from wheat (Fejes *et al.*, *Plant Mol. Biol.* 15:921-932 (1990), herein incorporated by reference in its entirety), the promoter for the *CAB-1* gene from spinach (Lubberstedt *et al.*, *Plant Physiol.* 104:997-1006 (1994), herein incorporated by reference in its entirety), the promoter for the *cab1R* gene from

rice (Luan *et al.*, *Plant Cell*. 4:971-981 (1992), the entirety of which is herein incorporated by reference), the pyruvate, orthophosphate dikinase (PPDK) promoter from maize (Matsuoka *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 90: 9586-9590 (1993), herein incorporated by reference in its entirety), the promoter for the tobacco Lhcb1\*2 gene (Cerdan *et al.*, *Plant Mol. Biol.* 33:245-255 (1997), herein incorporated by reference in its entirety), the *Arabidopsis thaliana* SUC2 sucrose-H<sup>+</sup> symporter promoter (Truernit *et al.*, *Planta*. 196:564-570 (1995), herein incorporated by reference in its entirety) and the promoter for the thylakoid membrane proteins from spinach (psaD, psaF, psaE, PC, FNR, atpC, atpD, cab, rbcS). Other promoters for the chlorophyll a/b-binding proteins may also be utilized in the present invention, such as the promoters for Lhcb gene and PsbP gene from white mustard (*Sinapis alba*; Kretsch *et al.*, *Plant Mol. Biol.* 28:219-229 (1995), the entirety of which is herein incorporated by reference).

For the purpose of expression in sink tissues of the plant, such as the tuber of the potato plant, the fruit of tomato, or the seed of maize, wheat, rice and barley, it is preferred that the promoters utilized in the present invention have relatively high expression in these specific tissues. A number of promoters for genes with tuber-specific or -enhanced expression are known, including the class I patatin promoter (Bevan *et al.*, *EMBO J.* 8:1899-1906 (1986); Jefferson *et al.*, *Plant Mol. Biol.* 14:995-1006 (1990), both of which are herein incorporated by reference in its entirety), the promoter for the potato tuber ADPGPP genes, both the large and small subunits, the sucrose synthase promoter (Salanoubat and Belliard, *Gene*. 60:47-56 (1987), Salanoubat and Belliard, *Gene*. 84:181-185 (1989), both of which are incorporated by reference in their entirety), the promoter for the major tuber proteins including the 22 kd protein complexes and proteinase inhibitors (Hannapel, *Plant Physiol.* 101:703-704 (1993), herein incorporated by reference in its entirety), the promoter for the granule bound starch synthase gene (GBSS)

(Visser *et al.*, *Plant Mol. Biol.* 17:691-699 (1991), herein incorporated by reference in its entirety) and other class I and II patatins promoters (Koster-Topfer *et al.*, *Mol Gen Genet.* 219:390-396 (1989); Mignery *et al.*, *Gene.* 62:27-44 (1988), both of which are herein incorporated by reference in their entirety).

Other promoters can also be used to express a sucrose pathway protein or fragment thereof in specific tissues, such as seeds or fruits. The promoter for  $\beta$ -conglycinin (Chen *et al.*, *Dev. Genet.* 10: 112-122 (1989), herein incorporated by reference in its entirety) or other seed-specific promoters such as the napin and phaseolin promoters, can be used. The zeins are a group of storage proteins found in maize endosperm. Genomic clones for zein genes have been isolated (Pedersen *et al.*, *Cell* 29:1015-1026 (1982), herein incorporated by reference in its entirety) and the promoters from these clones, including the 15 kD, 16 kD, 19 kD, 22 kD, 27 kD and genes, could also be used. Other promoters known to function, for example, in maize include the promoters for the following genes: *waxy*, *Brittle*, *Shrunken 2*, Branching enzymes I and II, starch synthases, debranching enzymes, oleosins, glutelins and sucrose synthases. A particularly preferred promoter for maize endosperm expression is the promoter for the glutelin gene from rice, more particularly the Osgt-1 promoter (Zheng *et al.*, *Mol. Cell Biol.* 13:5829-5842 (1993), herein incorporated by reference in its entirety). Examples of promoters suitable for expression in wheat include those promoters for the ADPglucose pyrosynthase (ADPGPP) subunits, the granule bound and other starch synthase, the branching and debranching enzymes, the embryogenesis-abundant proteins, the gliadins and the glutenins. Examples of such promoters in rice include those promoters for the ADPGPP subunits, the granule bound and other starch synthase, the branching enzymes, the debranching enzymes, sucrose synthases and the

glutelins. A particularly preferred promoter is the promoter for rice glutelin, Osgt-1. Examples of such promoters for barley include those for the ADPGPP subunits, the granule bound and other starch synthase, the branching enzymes, the debranching enzymes, sucrose synthases, the hordeins, the embryo globulins and the aleurone specific proteins.

Root specific promoters may also be used. An example of such a promoter is the promoter for the acid chitinase gene (Samac *et al.*, *Plant Mol. Biol.* 25:587-596 (1994), the entirety of which is herein incorporated by reference). Expression in root tissue could also be accomplished by utilizing the root specific subdomains of the CaMV35S promoter that have been identified (Lam *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 86:7890-7894 (1989), herein incorporated by reference in its entirety). Other root cell specific promoters include those reported by Conkling *et al.* (Conkling *et al.*, *Plant Physiol.* 93:1203-1211 (1990), the entirety of which is herein incorporated by reference).

Additional promoters that may be utilized are described, for example, in U.S. Patent Nos. 5,378,619; 5,391,725; 5,428,147; 5,447,858; 5,608,144; 5,608,144; 5,614,399; 5,633,441; 5,633,435; and 4,633,436, all of which are herein incorporated in their entirety. In addition, a tissue specific enhancer may be used (Fromm *et al.*, *The Plant Cell* 1:977-984 (1989), the entirety of which is herein incorporated by reference).

Constructs or vectors may also include with the coding region of interest a nucleic acid sequence that acts, in whole or in part, to terminate transcription of that region. For example, such sequences have been isolated including the Tr7 3' sequence and the NOS 3' sequence (Ingelbrecht *et al.*, *The Plant Cell* 1:671-680 (1989), the entirety of which is herein incorporated by reference; Bevan *et al.*, *Nucleic Acids Res.* 11:369-385 (1983), the entirety of which is herein incorporated by reference), or the like.



A vector or construct may also include regulatory elements. Examples of such include the Adh intron 1 (Callis *et al.*, *Genes and Develop. 1*:1183-1200 (1987), the entirety of which is herein incorporated by reference), the sucrose synthase intron (Vasil *et al.*, *Plant Physiol.* 91:1575-1579 (1989), the entirety of which is herein incorporated by reference) and the TMV omega element (Gallie *et al.*, *The Plant Cell 1*:301-311 (1989), the entirety of which is herein incorporated by reference). These and other regulatory elements may be included when appropriate.

A vector or construct may also include a selectable marker. Selectable markers may also be used to select for plants or plant cells that contain the exogenous genetic material. Examples of such include, but are not limited to, a neo gene (Potrykus *et al.*, *Mol. Gen. Genet.* 199:183-188 (1985), the entirety of which is herein incorporated by reference) which codes for kanamycin resistance and can be selected for using kanamycin, G418, etc.; a bar gene which codes for bialaphos resistance; a mutant EPSP synthase gene (Hinchee *et al.*, *Bio/Technology* 6:915-922 (1988), the entirety of which is herein incorporated by reference) which encodes glyphosate resistance; a nitrilase gene which confers resistance to bromoxynil (Stalker *et al.*, *J. Biol. Chem.* 263:6310-6314 (1988), the entirety of which is herein incorporated by reference); a mutant acetolactate synthase gene (ALS) which confers imidazolinone or sulphonylurea resistance (European Patent Application 154,204 (Sept. 11, 1985), the entirety of which is herein incorporated by reference); and a methotrexate resistant DHFR gene (Thillet *et al.*, *J. Biol. Chem.* 263:12500-12508 (1988), the entirety of which is herein incorporated by reference).

A vector or construct may also include a transit peptide. Incorporation of a suitable chloroplast transit peptide may also be employed (European Patent Application Publication Number 0218571, the entirety of which is herein incorporated by reference). Translational

enhancers may also be incorporated as part of the vector DNA. DNA constructs could contain one or more 5' non-translated leader sequences which may serve to enhance expression of the gene products from the resulting mRNA transcripts. Such sequences may be derived from the promoter selected to express the gene or can be specifically modified to increase translation of the mRNA. Such regions may also be obtained from viral RNAs, from suitable eukaryotic genes, or from a synthetic gene sequence. For a review of optimizing expression of transgenes, see Koziel *et al.*, *Plant Mol. Biol.* 32:393-405 (1996), the entirety of which is herein incorporated by reference.

A vector or construct may also include a screenable marker. Screenable markers may be used to monitor expression. Exemplary screenable markers include a  $\beta$ -glucuronidase or uidA gene (GUS) which encodes an enzyme for which various chromogenic substrates are known (Jefferson, *Plant Mol. Biol. Rep.* 5:387-405 (1987), the entirety of which is herein incorporated by reference; Jefferson *et al.*, *EMBO J.* 6:3901-3907 (1987), the entirety of which is herein incorporated by reference); an R-locus gene, which encodes a product that regulates the production of anthocyanin pigments (red color) in plant tissues (Dellaporta *et al.*, *Stadler Symposium* 11:263-282 (1988), the entirety of which is herein incorporated by reference); a  $\beta$ -lactamase gene (Sutcliffe *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 75:3737-3741 (1978), the entirety of which is herein incorporated by reference), a gene which encodes an enzyme for which various chromogenic substrates are known (e.g., PADAC, a chromogenic cephalosporin); a luciferase gene (Ow *et al.*, *Science* 234:856-859 (1986), the entirety of which is herein incorporated by reference); a xylE gene (Zukowsky *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 80:1101-1105 (1983), the entirety of which is herein incorporated by reference) which encodes a

catechol dioxygenase that can convert chromogenic catechols; an  $\alpha$ -amylase gene (Ikata *et al.*, *Bio/Technol.* 8:241-242 (1990), the entirety of which is herein incorporated by reference); a tyrosinase gene (Katz *et al.*, *J. Gen. Microbiol.* 129:2703-2714 (1983), the entirety of which is herein incorporated by reference) which encodes an enzyme capable of oxidizing tyrosine to DOPA and dopaquinone which in turn condenses to melanin; an  $\alpha$ -galactosidase, which will turn a chromogenic  $\alpha$ -galactose substrate.

Included within the terms "selectable or screenable marker genes" are also genes which encode a secretable marker whose secretion can be detected as a means of identifying or selecting for transformed cells. Examples include markers which encode a secretable antigen that can be identified by antibody interaction, or even secretable enzymes which can be detected catalytically. Secretable proteins fall into a number of classes, including small, diffusible proteins which are detectable, (*e.g.*, by ELISA), small active enzymes which are detectable in extracellular solution (*e.g.*,  $\alpha$ -amylase,  $\beta$ -lactamase, phosphinothricin transferase), or proteins which are inserted or trapped in the cell wall (such as proteins which include a leader sequence such as that found in the expression unit of extension or tobacco PR-S). Other possible selectable and/or screenable marker genes will be apparent to those of skill in the art.

There are many methods for introducing transforming nucleic acid molecules into plant cells. Suitable methods are believed to include virtually any method by which nucleic acid molecules may be introduced into a cell, such as by *Agrobacterium* infection or direct delivery of nucleic acid molecules such as, for example, by PEG-mediated transformation, by electroporation or by acceleration of DNA coated particles, etc (Potrykus, *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 42:205-225 (1991), the entirety of which is herein incorporated by

reference; Vasil, *Plant Mol. Biol.* 25:925-937 (1994), the entirety of which is herein incorporated by reference). For example, electroporation has been used to transform maize protoplasts (Fromm *et al.*, *Nature* 312:791-793 (1986), the entirety of which is herein incorporated by reference).

Other vector systems suitable for introducing transforming DNA into a host plant cell include but are not limited to binary artificial chromosome (BIBAC) vectors (Hamilton *et al.*, *Gene* 200:107-116 (1997), the entirety of which is herein incorporated by reference); and transfection with RNA viral vectors (Della-Cioppa *et al.*, *Ann. N.Y. Acad. Sci.* (1996), 792 (Engineering Plants for Commercial Products and Applications), 57-61, the entirety of which is herein incorporated by reference). Additional vector systems also include plant selectable YAC vectors such as those described in Mullen *et al.*, *Molecular Breeding* 4:449-457 (1988), the entirety of which is herein incorporated by reference).

Technology for introduction of DNA into cells is well known to those of skill in the art. Four general methods for delivering a gene into cells have been described: (1) chemical methods (Graham and van der Eb, *Virology* 54:536-539 (1973), the entirety of which is herein incorporated by reference); (2) physical methods such as microinjection (Capecchi, *Cell* 22:479-488 (1980), the entirety of which is herein incorporated by reference), electroporation (Wong and Neumann, *Biochem. Biophys. Res. Commun.* 107:584-587 (1982); Fromm *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 82:5824-5828 (1985); U.S. Patent No. 5,384,253, all of which are herein incorporated in their entirety); and the gene gun (Johnston and Tang, *Methods Cell Biol.* 43:353-365 (1994), the entirety of which is herein incorporated by reference); (3) viral vectors (Clapp, *Clin. Perinatol.* 20:155-168 (1993); Lu *et al.*, *J. Exp. Med.* 178:2089-2096 (1993); Eglitis and Anderson, *Biotechniques* 6:608-614 (1988), all of which are herein incorporated in their

entirety); and (4) receptor-mediated mechanisms (Curiel *et al.*, *Hum. Gen. Ther.* 3:147-154 (1992), Wagner *et al.*, *Proc. Natl. Acad. Sci. (USA)* 89:6099-6103 (1992), both of which are incorporated by reference in their entirety).

Acceleration methods that may be used include, for example, microprojectile bombardment and the like. One example of a method for delivering transforming nucleic acid molecules to plant cells is microprojectile bombardment. This method has been reviewed by Yang and Christou (eds.), *Particle Bombardment Technology for Gene Transfer*, Oxford Press, Oxford, England (1994), the entirety of which is herein incorporated by reference). Non-biological particles (microprojectiles) that may be coated with nucleic acids and delivered into cells by a propelling force. Exemplary particles include those comprised of tungsten, gold, platinum and the like.

A particular advantage of microprojectile bombardment, in addition to it being an effective means of reproducibly transforming monocots, is that neither the isolation of protoplasts (Cristou *et al.*, *Plant Physiol.* 87:671-674 (1988), the entirety of which is herein incorporated by reference) nor the susceptibility of *Agrobacterium* infection are required. An illustrative embodiment of a method for delivering DNA into maize cells by acceleration is a biolistics  $\alpha$ -particle delivery system, which can be used to propel particles coated with DNA through a screen, such as a stainless steel or Nytex screen, onto a filter surface covered with corn cells cultured in suspension. Gordon-Kamm *et al.*, describes the basic procedure for coating tungsten particles with DNA (Gordon-Kamm *et al.*, *Plant Cell* 2:603-618 (1990), the entirety of which is herein incorporated by reference). The screen disperses the tungsten nucleic acid particles so that they are not delivered to the recipient cells in large aggregates. A particle

delivery system suitable for use with the present invention is the helium acceleration PDS-1000/He gun is available from Bio-Rad Laboratories (Bio-Rad, Hercules, California)(Sanford *et al.*, *Technique* 3:3-16 (1991), the entirety of which is herein incorporated by reference).

For the bombardment, cells in suspension may be concentrated on filters. Filters containing the cells to be bombarded are positioned at an appropriate distance below the microprojectile stopping plate. If desired, one or more screens are also positioned between the gun and the cells to be bombarded.

Alternatively, immature embryos or other target cells may be arranged on solid culture medium. The cells to be bombarded are positioned at an appropriate distance below the microprojectile stopping plate. If desired, one or more screens are also positioned between the acceleration device and the cells to be bombarded. Through the use of techniques set forth herein one may obtain up to 1000 or more foci of cells transiently expressing a marker gene. The number of cells in a focus which express the exogenous gene product 48 hours post-bombardment often range from one to ten and average one to three.

In bombardment transformation, one may optimize the pre-bombardment culturing conditions and the bombardment parameters to yield the maximum numbers of stable transformants. Both the physical and biological parameters for bombardment are important in this technology. Physical factors are those that involve manipulating the DNA/microprojectile precipitate or those that affect the flight and velocity of either the macro- or microprojectiles. Biological factors include all steps involved in manipulation of cells before and immediately after bombardment, the osmotic adjustment of target cells to help alleviate the trauma associated with bombardment and also the nature of the transforming DNA, such as linearized DNA or

intact supercoiled plasmids. It is believed that pre-bombardment manipulations are especially important for successful transformation of immature embryos.

In another alternative embodiment, plastids can be stably transformed. Methods disclosed for plastid transformation in higher plants include the particle gun delivery of DNA containing a selectable marker and targeting of the DNA to the plastid genome through homologous recombination (Svab *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 87:8526-8530 (1990); Svab and Maliga, *Proc. Natl. Acad. Sci. (U.S.A.)* 90:913-917 (1993); Staub and Maliga, *EMBO J.* 12:601-606 (1993); U.S. Patents 5, 451,513 and 5,545,818, all of which are herein incorporated by reference in their entirety).

Accordingly, it is contemplated that one may wish to adjust various aspects of the bombardment parameters in small scale studies to fully optimize the conditions. One may particularly wish to adjust physical parameters such as gap distance, flight distance, tissue distance and helium pressure. One may also minimize the trauma reduction factors by modifying conditions which influence the physiological state of the recipient cells and which may therefore influence transformation and integration efficiencies. For example, the osmotic state, tissue hydration and the subculture stage or cell cycle of the recipient cells may be adjusted for optimum transformation. The execution of other routine adjustments will be known to those of skill in the art in light of the present disclosure.

*Agrobacterium*-mediated transfer is a widely applicable system for introducing genes into plant cells because the DNA can be introduced into whole plant tissues, thereby bypassing the need for regeneration of an intact plant from a protoplast. The use of *Agrobacterium*-mediated plant integrating vectors to introduce DNA into plant cells is well known in the art. See, for example the methods described by Fraley *et al.*, *Bio/Technology* 3:629-635 (1985) and Rogers *et*

al., *Methods Enzymol.* 153:253-277 (1987), both of which are herein incorporated by reference in their entirety. Further, the integration of the Ti-DNA is a relatively precise process resulting in few rearrangements. The region of DNA to be transferred is defined by the border sequences and intervening DNA is usually inserted into the plant genome as described (Spielmann *et al.*, *Mol. Gen. Genet.* 205:34 (1986), the entirety of which is herein incorporated by reference).

Modern *Agrobacterium* transformation vectors are capable of replication in *E. coli* as well as *Agrobacterium*, allowing for convenient manipulations as described (Klee *et al.*, In: *Plant DNA Infectious Agents*, Hohn and Schell (eds.), Springer-Verlag, New York, pp. 179-203 (1985), the entirety of which is herein incorporated by reference. Moreover, technological advances in vectors for *Agrobacterium*-mediated gene transfer have improved the arrangement of genes and restriction sites in the vectors to facilitate construction of vectors capable of expressing various polypeptide coding genes. The vectors described have convenient multi-linker regions flanked by a promoter and a polyadenylation site for direct expression of inserted polypeptide coding genes and are suitable for present purposes (Rogers *et al.*, *Methods Enzymol.* 153:253-277 (1987)). In addition, *Agrobacterium* containing both armed and disarmed Ti genes can be used for the transformations. In those plant strains where *Agrobacterium*-mediated transformation is efficient, it is the method of choice because of the facile and defined nature of the gene transfer.

A transgenic plant formed using *Agrobacterium* transformation methods typically contains a single gene on one chromosome. Such transgenic plants can be referred to as being heterozygous for the added gene. More preferred is a transgenic plant that is homozygous for the added structural gene; *i.e.*, a transgenic plant that contains two added genes, one gene at the same locus on each chromosome of a chromosome pair. A homozygous transgenic plant can be obtained by sexually mating (selfing) an independent segregant transgenic plant that contains a



single added gene, germinating some of the seed produced and analyzing the resulting plants produced for the gene of interest.

It is also to be understood that two different transgenic plants can also be mated to produce offspring that contain two independently segregating added, exogenous genes. Selfing of appropriate progeny can produce plants that are homozygous for both added, exogenous genes that encode a polypeptide of interest. Back-crossing to a parental plant and out-crossing with a non-transgenic plant are also contemplated, as is vegetative propagation.

Transformation of plant protoplasts can be achieved using methods based on calcium phosphate precipitation, polyethylene glycol treatment, electroporation and combinations of these treatments (*See, for example, Potrykus et al., Mol. Gen. Genet.* 205:193-200 (1986); Lorz *et al., Mol. Gen. Genet.* 199:178 (1985); Fromm *et al., Nature* 319:791 (1986); Uchimiya *et al., Mol. Gen. Genet.* 204:204 (1986); Marcotte *et al., Nature* 335:454-457 (1988), all of which are herein incorporated by reference in their entirety).

Application of these systems to different plant strains depends upon the ability to regenerate that particular plant strain from protoplasts. Illustrative methods for the regeneration of cereals from protoplasts are described (Fujimura *et al., Plant Tissue Culture Letters* 2:74 (1985); Toriyama *et al., Theor Appl. Genet.* 205:34 (1986); Yamada *et al., Plant Cell Rep.* 4:85 (1986); Abdullah *et al., Biotechnolog* 4:1087 (1986), all of which are herein incorporated by reference in their entirety).

To transform plant strains that cannot be successfully regenerated from protoplasts, other ways to introduce DNA into intact cells or tissues can be utilized. For example, regeneration of cereals from immature embryos or explants can be effected as described (Vasil, *Biotechnology* 6:397 (1988), the entirety of which is herein incorporated by reference). In addition, "particle

gun" or high-velocity microprojectile technology can be utilized (Vasil *et al.*, *Bio/Technology* 10:667 (1992), the entirety of which is herein incorporated by reference).

Using the latter technology, DNA is carried through the cell wall and into the cytoplasm on the surface of small metal particles as described (Klein *et al.*, *Nature* 328:70 (1987); Klein *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 85:8502-8505 (1988); McCabe *et al.*, *Bio/Technology* 6:923 (1988), all of which are herein incorporated by reference in their entirety). The metal particles penetrate through several layers of cells and thus allow the transformation of cells within tissue explants.

Other methods of cell transformation can also be used and include but are not limited to introduction of DNA into plants by direct DNA transfer into pollen (Zhou *et al.*, *Methods Enzymol.* 101:433 (1983); Hess *et al.*, *Intern Rev. Cytol.* 107:367 (1987); Luo *et al.*, *Plant Mol Biol. Reporter* 6:165 (1988), all of which are herein incorporated by reference in their entirety), by direct injection of DNA into reproductive organs of a plant (Pena *et al.*, *Nature* 325:274 (1987), the entirety of which is herein incorporated by reference), or by direct injection of DNA into the cells of immature embryos followed by the rehydration of desiccated embryos (Neuhaus *et al.*, *Theor. Appl. Genet.* 75:30 (1987), the entirety of which is herein incorporated by reference).

The regeneration, development and cultivation of plants from single plant protoplast transformants or from various transformed explants is well known in the art (Weissbach and Weissbach, In: *Methods for Plant Molecular Biology*, Academic Press, San Diego, CA, (1988), the entirety of which is herein incorporated by reference). This regeneration and growth process typically includes the steps of selection of transformed cells, culturing those individualized cells through the usual stages of embryonic development through the rooted plantlet stage. Transgenic

embryos and seeds are similarly regenerated. The resulting transgenic rooted shoots are thereafter planted in an appropriate plant growth medium such as soil.

The development or regeneration of plants containing the foreign, exogenous gene that encodes a protein of interest is well known in the art. Preferably, the regenerated plants are self-pollinated to provide homozygous transgenic plants. Otherwise, pollen obtained from the regenerated plants is crossed to seed-grown plants of agronomically important lines. Conversely, pollen from plants of these important lines is used to pollinate regenerated plants. A transgenic plant of the present invention containing a desired polypeptide is cultivated using methods well known to one skilled in the art.

There are a variety of methods for the regeneration of plants from plant tissue. The particular method of regeneration will depend on the starting plant tissue and the particular plant species to be regenerated.

Methods for transforming dicots, primarily by use of *Agrobacterium tumefaciens* and obtaining transgenic plants have been published for cotton (U.S. Patent No. 5,004,863; U.S. Patent No. 5,159,135; U.S. Patent No. 5,518,908, all of which are herein incorporated by reference in their entirety); soybean (U.S. Patent No. 5,569,834; U.S. Patent No. 5,416,011; McCabe *et al.*, *Biotechnology* 6:923 (1988); Christou *et al.*, *Plant Physiol.* 87:671-674 (1988); all of which are herein incorporated by reference in their entirety); *Brassica* (U.S. Patent No. 5,463,174, the entirety of which is herein incorporated by reference); peanut (Cheng *et al.*, *Plant Cell Rep.* 15:653-657 (1996), McKently *et al.*, *Plant Cell Rep.* 14:699-703 (1995), all of which are herein incorporated by reference in their entirety); papaya; and pea (Grant *et al.*, *Plant Cell Rep.* 15:254-258 (1995), the entirety of which is herein incorporated by reference).

Transformation of monocotyledons using electroporation, particle bombardment and *Agrobacterium* have also been reported. Transformation and plant regeneration have been achieved in asparagus (Bytebier *et al.*, *Proc. Natl. Acad. Sci. (USA)* 84:5354 (1987), the entirety of which is herein incorporated by reference); barley (Wan and Lemaux, *Plant Physiol* 104:37 (1994), the entirety of which is herein incorporated by reference); maize (Rhodes *et al.*, *Science* 240:204 (1988); Gordon-Kamm *et al.*, *Plant Cell* 2:603-618 (1990); Fromm *et al.*, *Bio/Technology* 8:833 (1990); Koziel *et al.*, *Bio/Technology* 11:194 (1993); Armstrong *et al.*, *Crop Science* 35:550-557 (1995); all of which are herein incorporated by reference in their entirety); oat (Somers *et al.*, *Bio/Technology* 10:1589 (1992), the entirety of which is herein incorporated by reference); orchard grass (Horn *et al.*, *Plant Cell Rep.* 7:469 (1988), the entirety of which is herein incorporated by reference); rice (Toriyama *et al.*, *Theor Appl. Genet.* 205:34 (1986); Part *et al.*, *Plant Mol. Biol.* 32:1135-1148 (1996); Abedinia *et al.*, *Aust. J. Plant Physiol.* 24:133-141 (1997); Zhang and Wu, *Theor. Appl. Genet.* 76:835 (1988); Zhang *et al.*, *Plant Cell Rep.* 7:379 (1988); Battraw and Hall, *Plant Sci.* 86:191-202 (1992); Christou *et al.*, *Bio/Technology* 9:957 (1991), all of which are herein incorporated by reference in their entirety); rye (De la Pena *et al.*, *Nature* 325:274 (1987), the entirety of which is herein incorporated by reference); sugarcane (Bower and Birch, *Plant J.* 2:409 (1992), the entirety of which is herein incorporated by reference); tall fescue (Wang *et al.*, *Bio/Technology* 10:691 (1992), the entirety of which is herein incorporated by reference) and wheat (Vasil *et al.*, *Bio/Technology* 10:667 (1992), the entirety of which is herein incorporated by reference); U.S. Patent No. 5,631,152, the entirety of which is herein incorporated by reference.)

Assays for gene expression based on the transient expression of cloned nucleic acid constructs have been developed by introducing the nucleic acid molecules into plant cells by

polyethylene glycol treatment, electroporation, or particle bombardment (Marcotte *et al.*, *Nature* 335:454-457 (1988), the entirety of which is herein incorporated by reference; Marcotte *et al.*, *Plant Cell* 1:523-532 (1989), the entirety of which is herein incorporated by reference; McCarty *et al.*, *Cell* 66:895-905 (1991), the entirety of which is herein incorporated by reference; Hattori *et al.*, *Genes Dev.* 6:609-618 (1992), the entirety of which is herein incorporated by reference; Goff *et al.*, *EMBO J.* 9:2517-2522 (1990), the entirety of which is herein incorporated by reference). Transient expression systems may be used to functionally dissect gene constructs (see generally, Mailga *et al.*, *Methods in Plant Molecular Biology*, Cold Spring Harbor Press (1995)).

Any of the nucleic acid molecules of the present invention may be introduced into a plant cell in a permanent or transient manner in combination with other genetic elements such as vectors, promoters, enhancers etc. Further, any of the nucleic acid molecules of the present invention may be introduced into a plant cell in a manner that allows for overexpression of the protein or fragment thereof encoded by the nucleic acid molecule.

Cosuppression is the reduction in expression levels, usually at the level of RNA, of a particular endogenous gene or gene family by the expression of a homologous sense construct that is capable of transcribing mRNA of the same strandedness as the transcript of the endogenous gene (Napoli *et al.*, *Plant Cell* 2:279-289 (1990), the entirety of which is herein incorporated by reference; van der Krol *et al.*, *Plant Cell* 2:291-299 (1990), the entirety of which is herein incorporated by reference). Cosuppression may result from stable transformation with a single copy nucleic acid molecule that is homologous to a nucleic acid sequence found with the cell (Proll and Meyer, *Plant J.* 2:465-475 (1992), the entirety of which is herein incorporated by reference) or with multiple copies of a nucleic acid molecule that is homologous to a nucleic acid

sequence found with the cell (Mittlesten *et al.*, *Mol. Gen. Genet.* 244:325-330 (1994), the entirety of which is herein incorporated by reference). Genes, even though different, linked to homologous promoters may result in the cosuppression of the linked genes (Vaucheret, *C.R. Acad. Sci. III* 316:1471-1483 (1993), the entirety of which is herein incorporated by reference).

This technique has, for example, been applied to generate white flowers from red petunia and tomatoes that do not ripen on the vine. Up to 50% of petunia transformants that contained a sense copy of the glucoamylase (CHS) gene produced white flowers or floral sectors; this was as a result of the post-transcriptional loss of mRNA encoding CHS (Flavell, *Proc. Natl. Acad. Sci. (U.S.A.)* 91:3490-3496 (1994), the entirety of which is herein incorporated by reference); van Blokland *et al.*, *Plant J.* 6:861-877 (1994), the entirety of which is herein incorporated by reference). Cosuppression may require the coordinate transcription of the transgene and the endogenous gene and can be reset by a developmental control mechanism (Jorgensen, *Trends Biotechnol.* 8:340-344 (1990), the entirety of which is herein incorporated by reference; Meins and Kunz, In: *Gene Inactivation and Homologous Recombination in Plants*, Paszkowski (ed.), pp. 335-348, Kluwer Academic, Netherlands (1994), the entirety of which is herein incorporated by reference).

It is understood that one or more of the nucleic acids of the present invention may be introduced into a plant cell and transcribed using an appropriate promoter with such transcription resulting in the cosuppression of an endogenous sucrose pathway protein.

Antisense approaches are a way of preventing or reducing gene function by targeting the genetic material (Mol *et al.*, *FEBS Lett.* 268:427-430 (1990), the entirety of which is herein incorporated by reference). The objective of the antisense approach is to use a sequence complementary to the target gene to block its expression and create a mutant cell line or

organism in which the level of a single chosen protein is selectively reduced or abolished.

Antisense techniques have several advantages over other 'reverse genetic' approaches. The site of inactivation and its developmental effect can be manipulated by the choice of promoter for antisense genes or by the timing of external application or microinjection. Antisense can manipulate its specificity by selecting either unique regions of the target gene or regions where it shares homology to other related genes (Hiatt *et al.*, In: *Genetic Engineering*, Setlow (ed.), Vol. 11, New York: Plenum 49-63 (1989), the entirety of which is herein incorporated by reference).

The principle of regulation by antisense RNA is that RNA that is complementary to the target mRNA is introduced into cells, resulting in specific RNA:RNA duplexes being formed by base pairing between the antisense substrate and the target mRNA (Green *et al.*, *Annu. Rev. Biochem.* 55:569-597 (1986), the entirety of which is herein incorporated by reference). Under one embodiment, the process involves the introduction and expression of an antisense gene sequence. Such a sequence is one in which part or all of the normal gene sequences are placed under a promoter in inverted orientation so that the 'wrong' or complementary strand is transcribed into a noncoding antisense RNA that hybridizes with the target mRNA and interferes with its expression (Takayama and Inouye, *Crit. Rev. Biochem. Mol. Biol.* 25:155-184 (1990), the entirety of which is herein incorporated by reference). An antisense vector is constructed by standard procedures and introduced into cells by transformation, transfection, electroporation, microinjection, infection, etc. The type of transformation and choice of vector will determine whether expression is transient or stable. The promoter used for the antisense gene may influence the level, timing, tissue, specificity, or inducibility of the antisense inhibition.

It is understood that the activity of a sucrose pathway protein in a plant cell may be reduced or depressed by growing a transformed plant cell containing a nucleic acid molecule whose non-transcribed strand encodes a sucrose pathway protein or fragment thereof.

Antibodies have been expressed in plants (Hiatt *et al.*, *Nature* 342:76-78 (1989), the entirety of which is herein incorporated by reference; Conrad and Fielder, *Plant Mol. Biol.* 26:1023-1030 (1994), the entirety of which is herein incorporated by reference). Cytoplasmic expression of a scFv (single-chain Fv antibodies) has been reported to delay infection by artichoke mottled crinkle virus. Transgenic plants that express antibodies directed against endogenous proteins may exhibit a physiological effect (Philips *et al.*, *EMBO J.* 16:4489-4496 (1997), the entirety of which is herein incorporated by reference; Marion-Poll, *Trends in Plant Science* 2:447-448 (1997), the entirety of which is herein incorporated by reference). For example, expressed anti-abscisic antibodies have been reported to result in a general perturbation of seed development (Philips *et al.*, *EMBO J.* 16: 4489-4496 (1997)).

Antibodies that are catalytic may also be expressed in plants (abzymes). The principle behind abzymes is that since antibodies may be raised against many molecules, this recognition ability can be directed toward generating antibodies that bind transition states to force a chemical reaction forward (Persidas, *Nature Biotechnology* 15:1313-1315 (1997), the entirety of which is herein incorporated by reference; Baca *et al.*, *Ann. Rev. Biophys. Biomol. Struct.* 26:461-493 (1997), the entirety of which is herein incorporated by reference). The catalytic abilities of abzymes may be enhanced by site directed mutagenesis. Examples of abzymes are, for example, set forth in U.S. Patent No: 5,658,753; U.S. Patent No. 5,632,990; U.S. Patent No. 5,631,137; U.S. Patent 5,602,015; U.S. Patent No. 5,559,538; U.S. Patent No. 5,576,174; U.S. Patent No.



5,500,358; U.S. Patent 5,318,897; U.S. Patent No. 5,298,409; U.S. Patent No. 5,258,289 and U.S. Patent No. 5,194,585, all of which are herein incorporated in their entirety.

It is understood that any of the antibodies of the present invention may be expressed in plants and that such expression can result in a physiological effect. It is also understood that any of the expressed antibodies may be catalytic.

**(b) Fungal Constructs and Fungal Transformants**

The present invention also relates to a fungal recombinant vector comprising exogenous genetic material. The present invention also relates to a fungal cell comprising a fungal recombinant vector. The present invention also relates to methods for obtaining a recombinant fungal host cell comprising introducing into a fungal host cell exogenous genetic material.

Exogenous genetic material may be transferred into a fungal cell. In a preferred embodiment the exogenous genetic material includes a nucleic acid molecule of the present invention having a sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof or fragments of either or other nucleic acid molecule of the present invention. The fungal recombinant vector may be any vector which can be conveniently subjected to recombinant DNA procedures. The choice of a vector will typically depend on the compatibility of the vector with the fungal host cell into which the vector is to be introduced. The vector may be a linear or a closed circular plasmid. The vector system may be a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the fungal host.

The fungal vector may be an autonomously replicating vector, *i.e.*, a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, *e.g.*, a plasmid, an extrachromosomal element, a minichromosome, or an artificial

chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the fungal cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. For integration, the vector may rely on the nucleic acid sequence of the vector for stable integration of the vector into the genome by homologous or nonhomologous recombination. Alternatively, the vector may contain additional nucleic acid sequences for directing integration by homologous recombination into the genome of the fungal host. The additional nucleic acid sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, there should be preferably two nucleic acid sequences which individually contain a sufficient number of nucleic acids, preferably 400bp to 1500bp, more preferably 800bp to 1000bp, which are highly homologous with the corresponding target sequence to enhance the probability of homologous recombination. These nucleic acid sequences may be any sequence that is homologous with a target sequence in the genome of the fungal host cell and, furthermore, may be non-encoding or encoding sequences.

For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. Examples of origin of replications for use in a yeast host cell are the 2 micron origin of replication and the combination of CEN3 and ARS 1. Any origin of replication may be used which is compatible with the fungal host cell of choice.

The fungal vectors of the present invention preferably contain one or more selectable markers which permit easy selection of transformed cells. A selectable marker is a gene the product of which provides, for example biocide or viral resistance, resistance to heavy metals,

prototrophy to auxotrophs and the like. The selectable marker may be selected from the group including, but not limited to, *amdS* (acetamidase), *argB* (ornithine carbamoyltransferase), *bar* (phosphinothricin acetyltransferase), *hygB* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *pyrG* (orotidine-5'-phosphate decarboxylase) and *sC* (sulfate adenylyltransferase) and *trpC* (anthranilate synthase). Preferred for use in an *Aspergillus* cell are the *amdS* and *pyrG* markers of *Aspergillus nidulans* or *Aspergillus oryzae* and the *bar* marker of *Streptomyces hygroscopicus*. Furthermore, selection may be accomplished by co-transformation, *e.g.*, as described in WO 91/17243, the entirety of which is herein incorporated by reference. A nucleic acid sequence of the present invention may be operably linked to a suitable promoter sequence. The promoter sequence is a nucleic acid sequence which is recognized by the fungal host cell for expression of the nucleic acid sequence. The promoter sequence contains transcription and translation control sequences which mediate the expression of the protein or fragment thereof.

A promoter may be any nucleic acid sequence which shows transcriptional activity in the fungal host cell of choice and may be obtained from genes encoding polypeptides either homologous or heterologous to the host cell. Examples of suitable promoters for directing the transcription of a nucleic acid construct of the invention in a filamentous fungal host are promoters obtained from the genes encoding *Aspergillus oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (*glaA*), *Rhizomucor miehei* lipase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Aspergillus nidulans* acetamidase and hybrids thereof. In a yeast host, a useful promoter is the *Saccharomyces cerevisiae* enolase (*eno-1*) promoter. Particularly preferred promoters are the TAKA amylase, NA2-tpi (a hybrid of the promoters from the genes encoding

*Aspergillus niger* neutral alpha -amylase and *Aspergillus oryzae* triose phosphate isomerase) and glaA promoters.

A protein or fragment thereof encoding nucleic acid molecule of the present invention may also be operably linked to a terminator sequence at its 3' terminus. The terminator sequence may be native to the nucleic acid sequence encoding the protein or fragment thereof or may be obtained from foreign sources. Any terminator which is functional in the fungal host cell of choice may be used in the present invention, but particularly preferred terminators are obtained from the genes encoding *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* glucoamylase, *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* alpha-glucosidase and *Saccharomyces cerevisiae* enolase.

A protein or fragment thereof encoding nucleic acid molecule of the present invention may also be operably linked to a suitable leader sequence. A leader sequence is a nontranslated region of a mRNA which is important for translation by the fungal host. The leader sequence is operably linked to the 5' terminus of the nucleic acid sequence encoding the protein or fragment thereof. The leader sequence may be native to the nucleic acid sequence encoding the protein or fragment thereof or may be obtained from foreign sources. Any leader sequence which is functional in the fungal host cell of choice may be used in the present invention, but particularly preferred leaders are obtained from the genes encoding *Aspergillus oryzae* TAKA amylase and *Aspergillus oryzae* triose phosphate isomerase.

A polyadenylation sequence may also be operably linked to the 3' terminus of the nucleic acid sequence of the present invention. The polyadenylation sequence is a sequence which when transcribed is recognized by the fungal host to add polyadenosine residues to transcribed mRNA. The polyadenylation sequence may be native to the nucleic acid sequence encoding the protein or

fragment thereof or may be obtained from foreign sources. Any polyadenylation sequence which is functional in the fungal host of choice may be used in the present invention, but particularly preferred polyadenylation sequences are obtained from the genes encoding *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* glucoamylase, *Aspergillus nidulans* anthranilate synthase and *Aspergillus niger* alpha-glucosidase.

To avoid the necessity of disrupting the cell to obtain the protein or fragment thereof and to minimize the amount of possible degradation of the expressed protein or fragment thereof within the cell, it is preferred that expression of the protein or fragment thereof gives rise to a product secreted outside the cell. To this end, a protein or fragment thereof of the present invention may be linked to a signal peptide linked to the amino terminus of the protein or fragment thereof. A signal peptide is an amino acid sequence which permits the secretion of the protein or fragment thereof from the fungal host into the culture medium. The signal peptide may be native to the protein or fragment thereof of the invention or may be obtained from foreign sources. The 5' end of the coding sequence of the nucleic acid sequence of the present invention may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region which encodes the secreted protein or fragment thereof. Alternatively, the 5' end of the coding sequence may contain a signal peptide coding region which is foreign to that portion of the coding sequence which encodes the secreted protein or fragment thereof. The foreign signal peptide may be required where the coding sequence does not normally contain a signal peptide coding region. Alternatively, the foreign signal peptide may simply replace the natural signal peptide to obtain enhanced secretion of the desired protein or fragment thereof. The foreign signal peptide coding region may be obtained from a glucoamylase or an amylase gene from an *Aspergillus* species, a lipase or proteinase gene

from *Rhizomucor miehei*, the gene for the alpha-factor from *Saccharomyces cerevisiae*, or the calf preprochymosin gene. An effective signal peptide for fungal host cells is the *Aspergillus oryzae* TKA amylase signal, *Aspergillus niger* neutral amylase signal, the *Rhizomucor miehei* aspartic proteinase signal, the *Humicola lanuginosus* cellulase signal, or the *Rhizomucor miehei* lipase signal. However, any signal peptide capable of permitting secretion of the protein or fragment thereof in a fungal host of choice may be used in the present invention.

A protein or fragment thereof encoding nucleic acid molecule of the present invention may also be linked to a propeptide coding region. A propeptide is an amino acid sequence found at the amino terminus of a protein or proenzyme. Cleavage of the propeptide from the proprotein yields a mature biochemically active protein. The resulting polypeptide is known as a propolypeptide or proenzyme (or a zymogen in some cases). Propolypeptides are generally inactive and can be converted to mature active polypeptides by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide or proenzyme. The propeptide coding region may be native to the protein or fragment thereof or may be obtained from foreign sources. The foreign propeptide coding region may be obtained from the *Saccharomyces cerevisiae* alpha-factor gene or *Myceliophthora thermophila* laccase gene (WO 95/33836, the entirety of which is herein incorporated by reference).

The procedures used to ligate the elements described above to construct the recombinant expression vector of the present invention are well known to one skilled in the art (see, for example, Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual*, 2nd ed., Cold Spring Harbor, N.Y., (1989)).

The present invention also relates to recombinant fungal host cells produced by the methods of the present invention which are advantageously used with the recombinant vector of

the present invention. The cell is preferably transformed with a vector comprising a nucleic acid sequence of the invention followed by integration of the vector into the host chromosome. The choice of fungal host cells will to a large extent depend upon the gene encoding the protein or fragment thereof and its source. The fungal host cell may, for example, be a yeast cell or a filamentous fungal cell.

"Yeast" as used herein includes *Ascosporogenous* yeast (*Endomycetales*), *Basidiosporogenous* yeast and yeast belonging to the *Fungi Imperfecti* (*Blastomycetes*). The *Ascosporogenous* yeasts are divided into the families *Spermophthoraceae* and *Saccharomycetaceae*. The latter is comprised of four subfamilies, *Schizosaccharomycoideae* (for example, genus *Schizosaccharomyces*), *Nadsonioideae*, *Lipomycoideae* and *Saccharomycoideae* (for example, genera *Pichia*, *Kluyveromyces* and *Saccharomyces*). The *Basidiosporogenous* yeasts include the genera *Leucosporidim*, *Rhodosporidium*, *Sporidiobolus*, *Filobasidium* and *Filobasidiella*. Yeast belonging to the *Fungi Imperfecti* are divided into two families, *Sporobolomycetaceae* (for example, genera *Sorobolomyces* and *Bullera*) and *Cryptococcaceae* (for example, genus *Candida*). Since the classification of yeast may change in the future, for the purposes of this invention, yeast shall be defined as described in Biology and Activities of Yeast (Skinner *et al.*, *Soc. App. Bacteriol. Symposium Series* No. 9, (1980), the entirety of which is herein incorporated by reference). The biology of yeast and manipulation of yeast genetics are well known in the art (*see*, for example, *Biochemistry and Genetics of Yeast*, Bacil *et al.* (ed.), 2nd edition, 1987; *The Yeasts*, Rose and Harrison (eds.), 2nd ed., (1987); and *The Molecular Biology of the Yeast Saccharomyces*, Strathern *et al.* (eds.), (1981), all of which are herein incorporated by reference in their entirety).

"Fungi" as used herein includes the phyla *Ascomycota*, *Basidiomycota*, *Chytridiomycota* and *Zygomycota* (as defined by Hawksworth *et al.*, In: Ainsworth and Bisby's *Dictionary of The Fungi*, 8th edition, 1995, CAB International, University Press, Cambridge, UK; the entirety of which is herein incorporated by reference) as well as the *Oomycota* (as cited in Hawksworth *et al.*, In: Ainsworth and Bisby's *Dictionary of The Fungi*, 8th edition, 1995, CAB International, University Press, Cambridge, UK) and all mitosporic fungi (Hawksworth *et al.*, In: Ainsworth and Bisby's *Dictionary of The Fungi*, 8th edition, 1995, CAB International, University Press, Cambridge, UK). Representative groups of *Ascomycota* include, for example, *Neurospora*, *Eupenicillium* (= *Penicillium*), *Emericella* (= *Aspergillus*), *Eurotium* (= *Aspergillus*) and the true yeasts listed above. Examples of *Basidiomycota* include mushrooms, rusts and smuts. Representative groups of *Chytridiomycota* include, for example, *Allomyces*, *Blastocladiella*, *Coelomomyces* and aquatic fungi. Representative groups of *Oomycota* include, for example, *Saprolegniomycetous* aquatic fungi (water molds) such as *Achlya*. Examples of mitosporic fungi include *Aspergillus*, *Penicillium*, *Candida* and *Alternaria*. Representative groups of *Zygomycota* include, for example, *Rhizopus* and *Mucor*.

"Filamentous fungi" include all filamentous forms of the subdivision *Eumycota* and *Oomycota* (as defined by Hawksworth *et al.*, In: Ainsworth and Bisby's *Dictionary of The Fungi*, 8th edition, 1995, CAB International, University Press, Cambridge, UK). The filamentous fungi are characterized by a vegetative mycelium composed of chitin, cellulose, glucan, chitosan, mannan and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative.



In one embodiment, the fungal host cell is a yeast cell. In a preferred embodiment, the yeast host cell is a cell of the species of *Candida*, *Kluyveromyces*, *Saccharomyces*, *Schizosaccharomyces*, *Pichia* and *Yarrowia*. In a preferred embodiment, the yeast host cell is a *Saccharomyces cerevisiae* cell, a *Saccharomyces carlsbergensis*, *Saccharomyces diastaticus* cell, a *Saccharomyces douglasii* cell, a *Saccharomyces kluyveri* cell, a *Saccharomyces norbensis* cell, or a *Saccharomyces oviformis* cell. In another preferred embodiment, the yeast host cell is a *Kluyveromyces lactis* cell. In another preferred embodiment, the yeast host cell is a *Yarrowia lipolytica* cell.

In another embodiment, the fungal host cell is a filamentous fungal cell. In a preferred embodiment, the filamentous fungal host cell is a cell of the species of, but not limited to, *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Myceliophthora*, *Mucor*, *Neurospora*, *Penicillium*, *Thielavia*, *Tolypocladium* and *Trichoderma*. In a preferred embodiment, the filamentous fungal host cell is an *Aspergillus* cell. In another preferred embodiment, the filamentous fungal host cell is an *Acremonium* cell. In another preferred embodiment, the filamentous fungal host cell is a *Fusarium* cell. In another preferred embodiment, the filamentous fungal host cell is a *Humicola* cell. In another preferred embodiment, the filamentous fungal host cell is a *Myceliophthora* cell. In another even preferred embodiment, the filamentous fungal host cell is a *Mucor* cell. In another preferred embodiment, the filamentous fungal host cell is a *Neurospora* cell. In another preferred embodiment, the filamentous fungal host cell is a *Penicillium* cell. In another preferred embodiment, the filamentous fungal host cell is a *Thielavia* cell. In another preferred embodiment, the filamentous fungal host cell is a *Tolypocladium* cell. In another preferred embodiment, the filamentous fungal host cell is a *Trichoderma* cell. In a preferred embodiment, the filamentous fungal host cell is an *Aspergillus*

*oryzae* cell, an *Aspergillus niger* cell, an *Aspergillus foetidus* cell, or an *Aspergillus japonicus* cell. In another preferred embodiment, the filamentous fungal host cell is a *Fusarium oxysporum* cell or a *Fusarium graminearum* cell. In another preferred embodiment, the filamentous fungal host cell is a *Humicola insolens* cell or a *Humicola lanuginosus* cell. In another preferred embodiment, the filamentous fungal host cell is a *Myceliophthora thermophila* cell. In a most preferred embodiment, the filamentous fungal host cell is a *Mucor miehei* cell. In a most preferred embodiment, the filamentous fungal host cell is a *Neurospora crassa* cell. In a most preferred embodiment, the filamentous fungal host cell is a *Penicillium purpurogenum* cell. In another most preferred embodiment, the filamentous fungal host cell is a *Thielavia terrestris* cell. In another most preferred embodiment, the *Trichoderma* cell is a *Trichoderma reesei* cell, a *Trichoderma viride* cell, a *Trichoderma longibrachiatum* cell, a *Trichoderma harzianum* cell, or a *Trichoderma koningii* cell. In a preferred embodiment, the fungal host cell is selected from an *A. nidulans* cell, an *A. niger* cell, an *A. oryzae* cell and an *A. sojae* cell. In a further preferred embodiment, the fungal host cell is an *A. nidulans* cell.

The recombinant fungal host cells of the present invention may further comprise one or more sequences which encode one or more factors that are advantageous in the expression of the protein or fragment thereof, for example, an activator (e.g., a trans-acting factor), a chaperone and a processing protease. The nucleic acids encoding one or more of these factors are preferably not operably linked to the nucleic acid encoding the protein or fragment thereof. An activator is a protein which activates transcription of a nucleic acid sequence encoding a polypeptide (Kudla *et al.*, *EMBO* 9:1355-1364(1990); Jarai and Buxton, *Current Genetics* 26:2238-244(1994); Verdier, *Yeast* 6:271-297(1990), all of which are herein incorporated by reference in their entirety). The nucleic acid sequence encoding an activator may be obtained

from the genes encoding *Saccharomyces cerevisiae* heme activator protein 1 (hap1), *Saccharomyces cerevisiae* galactose metabolizing protein 4 (gal4) and *Aspergillus nidulans* ammonia regulation protein (areA). For further examples, see Verdier, *Yeast* 6:271-297 (1990); MacKenzie *et al.*, *Journal of Gen. Microbiol.* 139:2295-2307 (1993), both of which are herein incorporated by reference in their entirety). A chaperone is a protein which assists another protein in folding properly (Hartl *et al.*, *TIBS* 19:20-25 (1994); Bergeron *et al.*, *TIBS* 19:124-128 (1994); Demolder *et al.*, *J. Biotechnology* 32:179-189 (1994); Craig, *Science* 260:1902-1903(1993); Gething and Sambrook, *Nature* 355:33-45 (1992); Puig and Gilbert, *J Biol. Chem.* 269:7764-7771 (1994); Wang and Tsou, *FASEB Journal* 7:1515-11157 (1993); Robinson *et al.*, *Bio/Technology* 1:381-384 (1994), all of which are herein incorporated by reference in their entirety). The nucleic acid sequence encoding a chaperone may be obtained from the genes encoding *Aspergillus oryzae* protein disulphide isomerase, *Saccharomyces cerevisiae* calnexin, *Saccharomyces cerevisiae* BiP/GRP78 and *Saccharomyces cerevisiae* Hsp70. For further examples, see Gething and Sambrook, *Nature* 355:33-45 (1992); Hartl *et al.*, *TIBS* 19:20-25 (1994). A processing protease is a protease that cleaves a propeptide to generate a mature biochemically active polypeptide (Enderlin and Ogrydziak, *Yeast* 10:67-79 (1994); Fuller *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 86:1434-1438 (1989); Julius *et al.*, *Cell* 37:1075-1089 (1984); Julius *et al.*, *Cell* 32:839-852 (1983), all of which are incorporated by reference in their entirety). The nucleic acid sequence encoding a processing protease may be obtained from the genes encoding *Aspergillus niger* Kex2, *Saccharomyces cerevisiae* dipeptidylaminopeptidase, *Saccharomyces cerevisiae* Kex2 and *Yarrowia lipolytica* dibasic processing endoprotease (xpr6). Any factor that is functional in the fungal host cell of choice may be used in the present invention.

66327-0100

Fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts and regeneration of the cell wall in a manner known per se. Suitable procedures for transformation of *Aspergillus* host cells are described in EP 238 023 and Yelton *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 81:1470-1474 (1984), both of which are herein incorporated by reference in their entirety. A suitable method of transforming *Fusarium* species is described by Malardier *et al.*, *Gene* 78:147-156 (1989), the entirety of which is herein incorporated by reference. Yeast may be transformed using the procedures described by Becker and Guarente, In: Abelson and Simon, (eds.), *Guide to Yeast Genetics and Molecular Biology, Methods Enzymol.* Volume 194, pp 182-187, Academic Press, Inc., New York; Ito *et al.*, *J. Bacteriology* 153:163 (1983); Hinnen *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 75:1920 (1978), all of which are herein incorporated by reference in their entirety.

The present invention also relates to methods of producing the protein or fragment thereof comprising culturing the recombinant fungal host cells under conditions conducive for expression of the protein or fragment thereof. The fungal cells of the present invention are cultivated in a nutrient medium suitable for production of the protein or fragment thereof using methods known in the art. For example, the cell may be cultivated by shake flask cultivation, small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the protein or fragment thereof to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art (*see, e.g.*, Bennett and LaSure (eds.), *More Gene Manipulations in Fungi*, Academic Press, CA, (1991), the entirety of which is herein incorporated by reference). Suitable media are available from commercial suppliers or may be

prepared according to published compositions (e.g., in catalogues of the American Type Culture Collection, Manassas, VA). If the protein or fragment thereof is secreted into the nutrient medium, a protein or fragment thereof can be recovered directly from the medium. If the protein or fragment thereof is not secreted, it is recovered from cell lysates.

The expressed protein or fragment thereof may be detected using methods known in the art that are specific for the particular protein or fragment. These detection methods may include the use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. For example, if the protein or fragment thereof has enzymatic activity, an enzyme assay may be used. Alternatively, if polyclonal or monoclonal antibodies specific to the protein or fragment thereof are available, immunoassays may be employed using the antibodies to the protein or fragment thereof. The techniques of enzyme assay and immunoassay are well known to those skilled in the art.

The resulting protein or fragment thereof may be recovered by methods known in the arts. For example, the protein or fragment thereof may be recovered from the nutrient medium by conventional procedures including, but not limited to, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation. The recovered protein or fragment thereof may then be further purified by a variety of chromatographic procedures, e.g., ion exchange chromatography, gel filtration chromatography, affinity chromatography, or the like.

### **(c) Mammalian Constructs and Transformed Mammalian Cells**

The present invention also relates to methods for obtaining a recombinant mammalian host cell, comprising introducing into a mammalian host cell exogenous genetic material. The present invention also relates to a mammalian cell comprising a mammalian recombinant vector. The present invention also relates to methods for obtaining a recombinant mammalian host cell,

comprising introducing into a mammalian cell exogenous genetic material. In a preferred embodiment the exogenous genetic material includes a nucleic acid molecule of the present invention having a sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof or fragments of either or other nucleic acid molecule of the present invention.

Mammalian cell lines available as hosts for expression are known in the art and include many immortalized cell lines available from the American Type Culture Collection (ATCC, Manassas, VA), such as HeLa cells, Chinese hamster ovary (CHO) cells, baby hamster kidney (BHK) cells and a number of other cell lines. Suitable promoters for mammalian cells are also known in the art and include viral promoters such as that from Simian Virus 40 (SV40) (Fiers *et al.*, *Nature* 273:113 (1978), the entirety of which is herein incorporated by reference), Rous sarcoma virus (RSV), adenovirus (ADV) and bovine papilloma virus (BPV). Mammalian cells may also require terminator sequences and poly-A addition sequences. Enhancer sequences which increase expression may also be included and sequences which promote amplification of the gene may also be desirable (for example methotrexate resistance genes).

Vectors suitable for replication in mammalian cells may include viral replicons, or sequences which insure integration of the appropriate sequences encoding HCV epitopes into the host genome. For example, another vector used to express foreign DNA is vaccinia virus. In this case, for example, a nucleic acid molecule encoding a protein or fragment thereof is inserted into the vaccinia genome. Techniques for the insertion of foreign DNA into the vaccinia virus genome are known in the art and may utilize, for example, homologous recombination. Such heterologous DNA is generally inserted into a gene which is non-essential to the virus, for example, the thymidine kinase gene (tk), which also provides a selectable marker. Plasmid

vectors that greatly facilitate the construction of recombinant viruses have been described (*see*, for example, Mackett *et al*, *J Virol.* 49:857 (1984); Chakrabarti *et al.*, *Mol. Cell. Biol.* 5:3403 (1985); Moss, In: *Gene Transfer Vectors For Mammalian Cells* (Miller and Calos, eds., Cold Spring Harbor Laboratory, N.Y., p. 10, (1987); all of which are herein incorporated by reference in their entirety). Expression of the HCV polypeptide then occurs in cells or animals which are infected with the live recombinant vaccinia virus.

The sequence to be integrated into the mammalian sequence may be introduced into the primary host by any convenient means, which includes calcium precipitated DNA, spheroplast fusion, transformation, electroporation, biolistics, lipofection, microinjection, or other convenient means. Where an amplifiable gene is being employed, the amplifiable gene may serve as the selection marker for selecting hosts into which the amplifiable gene has been introduced. Alternatively, one may include with the amplifiable gene another marker, such as a drug resistance marker, e.g. neomycin resistance (G418 in mammalian cells), hygromycin in resistance etc., or an auxotrophy marker (HIS3, TRP1, LEU2, URA3, ADE2, LYS2, etc.) for use in yeast cells.

Depending upon the nature of the modification and associated targeting construct, various techniques may be employed for identifying targeted integration. Conveniently, the DNA may be digested with one or more restriction enzymes and the fragments probed with an appropriate DNA fragment which will identify the properly sized restriction fragment associated with integration.

One may use different promoter sequences, enhancer sequences, or other sequence which will allow for enhanced levels of expression in the expression host. Thus, one may combine an enhancer from one source, a promoter region from another source, a 5'- noncoding region

upstream from the initiation sucrose from the same or different source as the other sequences and the like. One may provide for an intron in the non-coding region with appropriate splice sites or for an alternative 3'- untranslated sequence or polyadenylation site. Depending upon the particular purpose of the modification, any of these sequences may be introduced, as desired.

Where selection is intended, the sequence to be integrated will have with it a marker gene, which allows for selection. The marker gene may conveniently be downstream from the target gene and may include resistance to a cytotoxic agent, e.g. antibiotics, heavy metals, or the like, resistance or susceptibility to HAT, gancyclovir, etc., complementation to an auxotrophic host, particularly by using an auxotrophic yeast as the host for the subject manipulations, or the like. The marker gene may also be on a separate DNA molecule, particularly with primary mammalian cells. Alternatively, one may screen the various transformants, due to the high efficiency of recombination in yeast, by using hybridization analysis, PCR, sequencing, or the like.

For homologous recombination, constructs can be prepared where the amplifiable gene will be flanked, normally on both sides with DNA homologous with the DNA of the target region. Depending upon the nature of the integrating DNA and the purpose of the integration, the homologous DNA will generally be within 100kb, usually 50kb, preferably about 25kb, of the transcribed region of the target gene, more preferably within 2kb of the target gene. Where modeling of the gene is intended, homology will usually be present proximal to the site of the mutation. The homologous DNA may include the 5'-upstream region outside of the transcriptional regulatory region or comprising any enhancer sequences, transcriptional initiation sequences, adjacent sequences, or the like. The homologous region may include a portion of the coding region, where the coding region may be comprised only of an open reading frame or



combination of exons and introns. The homologous region may comprise all or a portion of an intron, where all or a portion of one or more exons may also be present. Alternatively, the homologous region may comprise the 3'-region, so as to comprise all or a portion of the transcriptional termination region, or the region 3' of this region. The homologous regions may extend over all or a portion of the target gene or be outside the target gene comprising all or a portion of the transcriptional regulatory regions and/or the structural gene.

The integrating constructs may be prepared in accordance with conventional ways, where sequences may be synthesized, isolated from natural sources, manipulated, cloned, ligated, subjected to in vitro mutagenesis, primer repair, or the like. At various stages, the joined sequences may be cloned and analyzed by restriction analysis, sequencing, or the like. Usually during the preparation of a construct where various fragments are joined, the fragments, intermediate constructs and constructs will be carried on a cloning vector comprising a replication system functional in a prokaryotic host, e.g., *E. coli* and a marker for selection, e.g., biocide resistance, complementation to an auxotrophic host, etc. Other functional sequences may also be present, such as polylinkers, for ease of introduction and excision of the construct or portions thereof, or the like. A large number of cloning vectors are available such as pBR322, the pUC series, etc. These constructs may then be used for integration into the primary mammalian host.

In the case of the primary mammalian host, a replicating vector may be used. Usually, such vector will have a viral replication system, such as SV40, bovine papilloma virus, adenovirus, or the like. The linear DNA sequence vector may also have a selectable marker for identifying transfected cells. Selectable markers include the neo gene, allowing for selection

with G418, the herpes tk gene for selection with HAT medium, the gpt gene with mycophenolic acid, complementation of an auxotrophic host, etc.

The vector may or may not be capable of stable maintenance in the host. Where the vector is capable of stable maintenance, the cells will be screened for homologous integration of the vector into the genome of the host, where various techniques for curing the cells may be employed. Where the vector is not capable of stable maintenance, for example, where a temperature sensitive replication system is employed, one may change the temperature from the permissive temperature to the non-permissive temperature, so that the cells may be cured of the vector. In this case, only those cells having integration of the construct comprising the amplifiable gene and, when present, the selectable marker, will be able to survive selection.

Where a selectable marker is present, one may select for the presence of the targeting construct by means of the selectable marker. Where the selectable marker is not present, one may select for the presence of the construct by the amplifiable gene. For the neo gene or the herpes tk gene, one could employ a medium for growth of the transformants of about 0.1-1 mg/ml of G418 or may use HAT medium, respectively. Where DHFR is the amplifiable gene, the selective medium may include from about 0.01-0.5 M of methotrexate or be deficient in glycine-hypoxanthine-thymidine and have dialysed serum (GHT media).

The DNA can be introduced into the expression host by a variety of techniques that include calcium phosphate/DNA co-precipitates, microinjection of DNA into the nucleus, electroporation, yeast protoplast fusion with intact cells, transfection, polycations, e.g., polybrene, polyornithine, etc., or the like. The DNA may be single or double stranded DNA, linear or circular. The various techniques for transforming mammalian cells are well known (see Keown *et al.*, *Methods Enzymol.* (1989); Keown *et al.*, *Methods Enzymol.* 185:527-537 (1990);

Mansour *et al.*, *Nature* 336:348-352, (1988); all of which are herein incorporated by reference in their entirety).

**(d) Insect Constructs and Transformed Insect Cells**

The present invention also relates to an insect recombinant vectors comprising exogenous genetic material. The present invention also relates to an insect cell comprising an insect recombinant vector. The present invention also relates to methods for obtaining a recombinant insect host cell, comprising introducing into an insect cell exogenous genetic material. In a preferred embodiment the exogenous genetic material includes a nucleic acid molecule of the present invention having a sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof or fragments of either or other nucleic acid molecule of the present invention.

The insect recombinant vector may be any vector which can be conveniently subjected to recombinant DNA procedures and can bring about the expression of the nucleic acid sequence. The choice of a vector will typically depend on the compatibility of the vector with the insect host cell into which the vector is to be introduced. The vector may be a linear or a closed circular plasmid. The vector system may be a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the insect host. In addition, the insect vector may be an expression vector. Nucleic acid molecules can be suitably inserted into a replication vector for expression in the insect cell under a suitable promoter for insect cells. Many vectors are available for this purpose and selection of the appropriate vector will depend mainly on the size of the nucleic acid molecule to be inserted into the vector and the particular host cell to be transformed with the vector. Each vector contains various components depending on its function (amplification of DNA or expression of DNA) and

the particular host cell with which it is compatible. The vector components for insect cell transformation generally include, but are not limited to, one or more of the following: a signal sequence, origin of replication, one or more marker genes and an inducible promoter.

The insect vector may be an autonomously replicating vector, *i.e.*, a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, *e.g.*, a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the insect cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. For integration, the vector may rely on the nucleic acid sequence of the vector for stable integration of the vector into the genome by homologous or nonhomologous recombination. Alternatively, the vector may contain additional nucleic acid sequences for directing integration by homologous recombination into the genome of the insect host. The additional nucleic acid sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, there should be preferably two nucleic acid sequences which individually contain a sufficient number of nucleic acids, preferably 400bp to 1500bp, more preferably 800bp to 1000bp, which are highly homologous with the corresponding target sequence to enhance the probability of homologous recombination. These nucleic acid sequences may be any sequence that is homologous with a target sequence in the genome of the insect host cell and, furthermore, may be non-encoding or encoding sequences.

Baculovirus expression vectors (BEVs) have become important tools for the expression of foreign genes, both for basic research and for the production of proteins with direct clinical

applications in human and veterinary medicine (Doerfler, *Curr. Top. Microbiol. Immunol.* 131:51-68 (1968); Luckow and Summers, *Bio/Technology* 6:47-55 (1988a); Miller, *Annual Review of Microbiol.* 42:177-199 (1988); Summers, *Curr. Comm. Molecular Biology*, Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (1988); all of which are herein incorporated by reference in their entirety). BEVs are recombinant insect viruses in which the coding sequence for a chosen foreign gene has been inserted behind a baculovirus promoter in place of the viral gene, e.g., polyhedrin (Smith and Summers, U.S. Pat. No., 4,745,051, the entirety of which is incorporated herein by reference).

The use of baculovirus vectors relies upon the host cells being derived from *Lepidopteran* insects such as *Spodoptera frugiperda* or *Trichoplusia ni*. The preferred *Spodoptera frugiperda* cell line is the cell line Sf9. The *Spodoptera frugiperda* Sf9 cell line was obtained from American Type Culture Collection (Manassas, VA.) and is assigned accession number ATCC CRL 1711 (Summers and Smith, *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures*, Texas Ag. Exper. Station Bulletin No. 1555 (1988), the entirety of which is herein incorporated by reference). Other insect cell systems, such as the silkworm *B. mori* may also be used.

The proteins expressed by the BEVs are, therefore, synthesized, modified and transported in host cells derived from *Lepidopteran* insects. Most of the genes that have been inserted and produced in the baculovirus expression vector system have been derived from vertebrate species. Other baculovirus genes in addition to the polyhedrin promoter may be employed to advantage in a baculovirus expression system. These include immediate-early (alpha), delayed-early ( ), late ( ), or very late (delta), according to the phase of the viral infection during which they are expressed. The expression of these genes occurs sequentially, probably as the result of a

"cascade" mechanism of transcriptional regulation. (Guarino and Summers, *J. Virol.* 57:563-571 (1986); Guarino and Summers, *J. Virol.* 61:2091-2099 (1987); Guarino and Summers, *Virol.* 162:444-451 (1988); all of which are herein incorporated by reference in their entirety).

Insect recombinant vectors are useful as intermediates for the infection or transformation of insect cell systems. For example, an insect recombinant vector containing a nucleic acid molecule encoding a baculovirus transcriptional promoter followed downstream by an insect signal DNA sequence is capable of directing the secretion of the desired biologically active protein from the insect cell. The vector may utilize a baculovirus transcriptional promoter region derived from any of the over 500 baculoviruses generally infecting insects, such as for example the Orders *Lepidoptera*, *Diptera*, *Orthoptera*, *Coleoptera* and *Hymenoptera*, including for example but not limited to the viral DNAs of *Autographa californica* MNPV, *Bombyx mori* NPV, *Trichoplusia ni* MNPV, *Rachiplusia ou* MNPV or *Galleria mellonella* MNPV, wherein said baculovirus transcriptional promoter is a baculovirus immediate-early gene IEl or IEN promoter; an immediate-early gene in combination with a baculovirus delayed-early gene promoter region selected from the group consisting of 39K and a *HindIII-k* fragment delayed-early gene; or a baculovirus late gene promoter. The immediate-early or delayed-early promoters can be enhanced with transcriptional enhancer elements. The insect signal DNA sequence may code for a signal peptide of a *Lepidopteran* adipokinetic hormone precursor or a signal peptide of the *Manduca sexta* adipokinetic hormone precursor (Summers, U.S. Patent No. 5,155,037; the entirety of which is herein incorporated by reference). Other insect signal DNA sequences include a signal peptide of the *Orthoptera Schistocerca gregaria* locust adipokinetic hormone precursor and the *Drosophila melanogaster* cuticle genes CP1, CP2, CP3 or CP4 or for an insect

signal peptide having substantially a similar chemical composition and function (Summers, U.S. Patent No. 5,155,037).

Insect cells are distinctly different from animal cells. Insects have a unique life cycle and have distinct cellular properties such as the lack of intracellular plasminogen activators in which are present in vertebrate cells. Another difference is the high expression levels of protein products ranging from 1 to greater than 500 mg/liter and the ease at which cDNA can be cloned into cells (Frasier, *In Vitro Cell. Dev. Biol.* 25:225 (1989); Summers and Smith, In: *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures*, Texas Ag. Exper. Station Bulletin No. 1555 (1988), both of which are incorporated by reference in their entirety).

Recombinant protein expression in insect cells is achieved by viral infection or stable transformation. For viral infection, the desired gene is cloned into baculovirus at the site of the wild-type polyhedron gene (Webb and Summers, *Technique* 2:173 (1990); Bishop and Posse, *Adv. Gene Technol.* 1:55 (1990); both of which are incorporated by reference in their entirety). The polyhedron gene is a component of a protein coat in occlusions which encapsulate virus particles. Deletion or insertion in the polyhedron gene results the failure to form occlusion bodies. Occlusion negative viruses are morphologically different from occlusion positive viruses and enable one skilled in the art to identify and purify recombinant viruses.

The vectors of present invention preferably contain one or more selectable markers which permit easy selection of transformed cells. A selectable marker is a gene the product of which provides, for example biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs and the like. Selection may be accomplished by co-transformation, *e.g.*, as described in WO 91/17243, a nucleic acid sequence of the present invention may be operably linked to a suitable promoter sequence. The promoter sequence is a nucleic acid sequence which is

recognized by the insect host cell for expression of the nucleic acid sequence. The promoter sequence contains transcription and translation control sequences which mediate the expression of the protein or fragment thereof. The promoter may be any nucleic acid sequence which shows transcriptional activity in the insect host cell of choice and may be obtained from genes encoding polypeptides either homologous or heterologous to the host cell.

For example, a nucleic acid molecule encoding a protein or fragment thereof may also be operably linked to a suitable leader sequence. A leader sequence is a nontranslated region of a mRNA which is important for translation by the fungal host. The leader sequence is operably linked to the 5' terminus of the nucleic acid sequence encoding the protein or fragment thereof. The leader sequence may be native to the nucleic acid sequence encoding the protein or fragment thereof or may be obtained from foreign sources. Any leader sequence which is functional in the insect host cell of choice may be used in the present invention.

A polyadenylation sequence may also be operably linked to the 3' terminus of the nucleic acid sequence of the present invention. The polyadenylation sequence is a sequence which when transcribed is recognized by the insect host to add polyadenosine residues to transcribed mRNA. The polyadenylation sequence may be native to the nucleic acid sequence encoding the protein or fragment thereof or may be obtained from foreign sources. Any polyadenylation sequence which is functional in the fungal host of choice may be used in the present invention.

To avoid the necessity of disrupting the cell to obtain the protein or fragment thereof and to minimize the amount of possible degradation of the expressed polypeptide within the cell, it is preferred that expression of the polypeptide gene gives rise to a product secreted outside the cell. To this end, the protein or fragment thereof of the present invention may be linked to a signal peptide linked to the amino terminus of the protein or fragment thereof. A signal peptide is an



amino acid sequence which permits the secretion of the protein or fragment thereof from the insect host into the culture medium. The signal peptide may be native to the protein or fragment thereof of the invention or may be obtained from foreign sources. The 5' end of the coding sequence of the nucleic acid sequence of the present invention may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region which encodes the secreted protein or fragment thereof.

At present, a mode of achieving secretion of a foreign gene product in insect cells is by way of the foreign gene's native signal peptide. Because the foreign genes are usually from non-insect organisms, their signal sequences may be poorly recognized by insect cells and hence, levels of expression may be suboptimal. However, the efficiency of expression of foreign gene products seems to depend primarily on the characteristics of the foreign protein. On average, nuclear localized or non-structural proteins are most highly expressed, secreted proteins are intermediate and integral membrane proteins are the least expressed. One factor generally affecting the efficiency of the production of foreign gene products in a heterologous host system is the presence of native signal sequences (also termed presequences, targeting signals, or leader sequences) associated with the foreign gene. The signal sequence is generally coded by a DNA sequence immediately following (5' to 3') the translation start site of the desired foreign gene.

The expression dependence on the type of signal sequence associated with a gene product can be represented by the following example: If a foreign gene is inserted at a site downstream from the translational start site of the baculovirus polyhedrin gene so as to produce a fusion protein (containing the N-terminus of the polyhedrin structural gene), the fused gene is highly expressed. But less expression is achieved when a foreign gene is inserted in a baculovirus

expression vector immediately following the transcriptional start site and totally replacing the polyhedrin structural gene.

Insertions into the region -50 to -1 significantly alter (reduce) steady state transcription which, in turn, reduces translation of the foreign gene product. Use of the pVL941 vector optimizes transcription of foreign genes to the level of the polyhedrin gene transcription. Even though the transcription of a foreign gene may be optimal, optimal translation may vary because of several factors involving processing: signal peptide recognition, mRNA and ribosome binding, glycosylation, disulfide bond formation, sugar processing, oligomerization, for example.

The properties of the insect signal peptide are expected to be more optimal for the efficiency of the translation process in insect cells than those from vertebrate proteins. This phenomenon can generally be explained by the fact that proteins secreted from cells are synthesized as precursor molecules containing hydrophobic N-terminal signal peptides. The signal peptides direct transport of the select protein to its target membrane and are then cleaved by a peptidase on the membrane, such as the endoplasmic reticulum, when the protein passes through it.

Another exemplary insect signal sequence is the sequence encoding for *Drosophila* cuticle proteins such as CP1, CP2, CP3 or CP4 (Summers, U.S. Patent No. 5,278,050; the entirety of which is herein incorporated by reference). Most of a 9kb region of the *Drosophila* genome containing genes for the cuticle proteins has been sequenced. Four of the five cuticle genes contains a signal peptide coding sequence interrupted by a short intervening sequence (about 60 base pairs) at a conserved site. Conserved sequences occur in the 5' mRNA untranslated region, in the adjacent 35 base pairs of upstream flanking sequence and at -200 base pairs from the mRNA start position in each of the cuticle genes.

Standard methods of insect cell culture, cotransfection and preparation of plasmids are set forth in Summers and Smith (Summers and Smith, *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures*, Texas Agricultural Experiment Station Bulletin No. 1555, Texas A&M University (1987)). Procedures for the cultivation of viruses and cells are described in Volkman and Summers, *J. Virol* 19:820-832 (1975) and Volkman *et al.*, *J. Virol* 19:820-832 (1976); both of which are herein incorporated by reference in their entirety.

**(e) Bacterial Constructs and Transformed Bacterial Cells**

The present invention also relates to a bacterial recombinant vector comprising exogenous genetic material. The present invention also relates to a bacteria cell comprising a bacterial recombinant vector. The present invention also relates to methods for obtaining a recombinant bacteria host cell, comprising introducing into a bacterial host cell exogenous genetic material. . In a preferred embodiment the exogenous genetic material includes a nucleic acid molecule of the present invention having a sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 2814 or complements thereof or fragments of either or other nucleic acid molecule of the present invention.

The bacterial recombinant vector may be any vector which can be conveniently subjected to recombinant DNA procedures. The choice of a vector will typically depend on the compatibility of the vector with the bacterial host cell into which the vector is to be introduced. The vector may be a linear or a closed circular plasmid. The vector system may be a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the bacterial host. In addition, the bacterial vector may be an expression vector. Nucleic acid molecules encoding protein homologues or fragments thereof can, for example, be suitably inserted into a replicable vector for expression in the bacterium

under the control of a suitable promoter for bacteria. Many vectors are available for this purpose and selection of the appropriate vector will depend mainly on the size of the nucleic acid to be inserted into the vector and the particular host cell to be transformed with the vector. Each vector contains various components depending on its function (amplification of DNA or expression of DNA) and the particular host cell with which it is compatible. The vector components for bacterial transformation generally include, but are not limited to, one or more of the following: a signal sequence, an origin of replication, one or more marker genes and an inducible promoter.

In general, plasmid vectors containing replicon and control sequences that are derived from species compatible with the host cell are used in connection with bacterial hosts. The vector ordinarily carries a replication site, as well as marking sequences that are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is typically transformed using pBR322, a plasmid derived from an *E. coli* species (see, e.g., Bolivar *et al.*, *Gene* 2:95 (1977); the entirety of which is herein incorporated by reference). pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid or phage, also generally contains, or is modified to contain, promoters that can be used by the microbial organism for expression of the selectable marker genes.

Nucleic acid molecules encoding protein or fragments thereof may be expressed not only directly, but also as a fusion with another polypeptide, preferably a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the polypeptide DNA that is inserted into the vector. The heterologous signal sequence selected

should be one that is recognized and processed (i.e., cleaved by a signal peptidase) by the host cell. For bacterial host cells that do not recognize and process the native polypeptide signal sequence, the signal sequence is substituted by a bacterial signal sequence selected, for example, from the group consisting of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Generally, in cloning vectors this sequence is one that enables the vector to replicate independently of the host chromosomal DNA and includes origins of replication or autonomously replicating sequences. Such sequences are well known for a variety of bacteria. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria.

Expression and cloning vectors also generally contain a selection gene, also termed a selectable marker. This gene encodes a protein necessary for the survival or growth of transformed host cells grown in a selective culture medium. Host cells not transformed with the vector containing the selection gene will not survive in the culture medium. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*. One example of a selection scheme utilizes a drug to arrest growth of a host cell. Those cells that are successfully transformed with a heterologous protein homologue or fragment thereof produce a protein conferring drug resistance and thus survive the selection regimen.

The expression vector for producing a protein or fragment thereof can also contain an inducible promoter that is recognized by the host bacterial organism and is operably linked to the

nucleic acid encoding, for example, the nucleic acid molecule encoding the protein homologue or fragment thereof of interest. Inducible promoters suitable for use with bacterial hosts include the -lactamase and lactose promoter systems (Chang *et al.*, *Nature* 275:615 (1978); Goeddel *et al.*, *Nature* 281:544 (1979); both of which are herein incorporated by reference in their entirety), the arabinose promoter system (Guzman *et al.*, *J. Bacteriol.* 174:7716-7728 (1992); the entirety of which is herein incorporated by reference), alkaline phosphatase, a tryptophan (trp) promoter system (Goeddel, *Nucleic Acids Res.* 8:4057 (1980); EP 36,776; both of which are herein incorporated by reference in their entirety) and hybrid promoters such as the tac promoter (deBoer *et al.*, *Proc. Natl. Acad. Sci. (USA)* 80:21-25 (1983); the entirety of which is herein incorporated by reference). However, other known bacterial inducible promoters are suitable (Siebenlist *et al.*, *Cell* 20:269 (1980); the entirety of which is herein incorporated by reference).

Promoters for use in bacterial systems also generally contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the polypeptide of interest. The promoter can be removed from the bacterial source DNA by restriction enzyme digestion and inserted into the vector containing the desired DNA.

Construction of suitable vectors containing one or more of the above-listed components employs standard ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored and re-ligated in the form desired to generate the plasmids required. Examples of available bacterial expression vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as Bluescript™ (Stratagene, La Jolla, CA), in which, for example, encoding an *A. nidulans* protein homologue or fragment thereof homologue, may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of -

galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke and Schuster, *J. Biol. Chem.* 264:5503-5509 (1989), the entirety of which is herein incorporated by reference); and the like. pGEX vectors (Promega, Madison Wisconsin U.S.A.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems are designed to include heparin, thrombin or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

Suitable host bacteria for a bacterial vector include archaeobacteria and eubacteria, especially eubacteria and most preferably *Enterobacteriaceae*. Examples of useful bacteria include *Escherichia*, *Enterobacter*, *Azotobacter*, *Erwinia*, *Bacillus*, *Pseudomonas*, *Klebsiella*, *Proteus*, *Salmonella*, *Serratia*, *Shigella*, *Rhizobia*, *Vitreoscilla* and *Paracoccus*. Suitable *E. coli* hosts include *E. coli* W3110 (American Type Culture Collection (ATCC) 27,325, Manassas, Virginia U.S.A.), *E. coli* 294 (ATCC 31,446), *E. coli* B and *E. coli* X1776 (ATCC 31,537). These examples are illustrative rather than limiting. Mutant cells of any of the above-mentioned bacteria may also be employed. It is, of course, necessary to select the appropriate bacteria taking into consideration replicability of the replicon in the cells of a bacterium. For example, *E. coli*, *Serratia*, or *Salmonella* species can be suitably used as the host when well known plasmids such as pBR322, pBR325, pACYC177, or pKN410 are used to supply the replicon. *E. coli* strain W3110 is a preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell should secrete minimal amounts of proteolytic enzymes.

Host cells are transfected and preferably transformed with the above-described vectors and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences.

Numerous methods of transfection are known to the ordinarily skilled artisan, for example, calcium phosphate and electroporation. Depending on the host cell used, transformation is done using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in section 1.82 of Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, New York: Cold Spring Harbor Laboratory Press, (1989), is generally used for bacterial cells that contain substantial cell-wall barriers. Another method for transformation employs polyethylene glycol/DMSO, as described in Chung and Miller (Chung and Miller, *Nucleic Acids Res.* 16:3580 (1988); the entirety of which is herein incorporated by reference). Yet another method is the use of the technique termed electroporation.

Bacterial cells used to produce the polypeptide of interest for purposes of this invention are cultured in suitable media in which the promoters for the nucleic acid encoding the heterologous polypeptide can be artificially induced as described generally, e.g., in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, New York: Cold Spring Harbor Laboratory Press, (1989). Examples of suitable media are given in U.S. Pat. Nos. 5,304,472 and 5,342,763; both of which are incorporated by reference in their entirety.

In addition to the above discussed procedures, practitioners are familiar with the standard resource materials which describe specific conditions and procedures for the construction, manipulation and isolation of macromolecules (e.g., DNA molecules, plasmids, etc.), generation of recombinant organisms and the screening and isolating of clones, (see for example, Sambrook



*et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (1989); Mailga *et al.*, *Methods in Plant Molecular Biology*, Cold Spring Harbor Press (1995), the entirety of which is herein incorporated by reference; Birren *et al.*, *Genome Analysis: Analyzing DNA*, 1, Cold Spring Harbor, New York, the entirety of which is herein incorporated by reference).

**(f) Computer Readable Media**

The nucleotide sequence provided in SEQ ID NO: 1 through SEQ ID NO: 2814 or fragment thereof, or complement thereof, or a nucleotide sequence at least 90% identical, preferably 95%, identical even more preferably 99% or 100% identical to the sequence provided in SEQ ID NO: 1 through SEQ ID NO: 2814 or fragment thereof, or complement thereof, can be “provided” in a variety of mediums to facilitate use. Such a medium can also provide a subset thereof in a form that allows a skilled artisan to examine the sequences.

A preferred subset of nucleotide sequences are those nucleic acid sequences that encodes a maize or a soybean triose phosphate isomerase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate aldolase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean fructose 1,6-bisphosphate enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean fructose 6-phosphate 2-kinase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean phosphoglucoisomerase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean vacuolar H<sup>+</sup> translocating-pyrophosphatase enzyme or complement thereof or fragment of either, a nucleic acid molecule that encodes a maize or a soybean pyrophosphate-dependent fructose-6-phosphate phosphotransferase enzyme or complement thereof or fragment of either, a nucleic acid molecule





be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention.

As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate media comprising the nucleotide sequence information of the present invention. A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing one or more of nucleotide sequences of the present invention, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990), the entirety of which is herein incorporated by reference) and BLAZE (Brutlag *et al.*, *Comp. Chem.*

17:203-207 (1993), the entirety of which is herein incorporated by reference) search algorithms on a Sybase system can be used to identify open reading frames (ORFs) within the genome that contain homology to ORFs or proteins from other organisms. Such ORFs are protein-encoding fragments within the sequences of the present invention and are useful in producing commercially important proteins such as enzymes used in amino acid biosynthesis, metabolism, transcription, translation, RNA processing, nucleic acid and a protein degradation, protein modification and DNA replication, restriction, modification, recombination and repair.

The present invention further provides systems, particularly computer-based systems, which contain the sequence information described herein. Such systems are designed to identify commercially important fragments of the nucleic acid molecule of the present invention. As used herein, "a computer-based system" refers to the hardware means, software means and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based system are suitable for use in the present invention.

As indicated above, the computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory that can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention. As used herein, "search means" refers to one or more programs which are implemented on the computer-

based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequence of the present invention that match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are available can be used in the computer-based systems of the present invention. Examples of such software include, but are not limited to, MacPattern (EMBL), BLASTIN and BLASTIX (NCBIA). One of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems.

The most preferred sequence length of a target sequence is from about 10 to 100 amino acids or from about 30 to 300 nucleotide residues. However, it is well recognized that during searches for commercially important fragments of the nucleic acid molecules of the present invention, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequences the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzymatic active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, *cis* elements, hairpin structures and inducible expression elements (protein binding sequences).

Thus, the present invention further provides an input means for receiving a target sequence, a data storage means for storing the target sequences of the present invention sequence

identified using a search means as described above and an output means for outputting the identified homologous sequences. A variety of structural formats for the input and output means can be used to input and output information in the computer-based systems of the present invention. A preferred format for an output means ranks fragments of the sequence of the present invention by varying degrees of homology to the target sequence or target motif. Such presentation provides a skilled artisan with a ranking of sequences which contain various amounts of the target sequence or target motif and identifies the degree of homology contained in the identified fragment.

A variety of comparing means can be used to compare a target sequence or target motif with the data storage means to identify sequence fragments sequence of the present invention. For example, implementing software which implement the BLAST and BLAZE algorithms (Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990)) can be used to identify open frames within the nucleic acid molecules of the present invention. A skilled artisan can readily recognize that any one of the publicly available homology search programs can be used as the search means for the computer-based systems of the present invention.

Having now generally described the invention, the same will be more readily understood through reference to the following examples which are provided by way of illustration and are not intended to be limiting of the present invention, unless specified.

### **Example 1**

The MONN01 cDNA library is a normalized library generated from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) total leaf tissue at the V6 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the

same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is at the 6-leaf development stage. The older, more juvenile leaves, which are in a basal position, as well as the younger, more adult leaves, which are more apical are cut at the base of the leaves. The leaves are then pooled and immediately transferred to liquid nitrogen containers in which the pooled leaves are crushed. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON001 cDNA library is generated from maize (B73, Illinois Foundation Seeds, Champaign, Illinois U.S.A.) immature tassels at the V6 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in a greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue from the maize plant is collected at the V6 stage. At that stage the



tassel is an immature tassel of about 2-3 cm in length. The tassels are removed and frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON003 library is generated from maize (B73 x Mo17, Illinois Foundation Seeds, Champaign, Illinois U.S.A.) roots at the V6 developmental stage. Seeds are planted at a depth of approximately 3 cm in coil into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth, the seedlings are transplanted into 10 inch pots containing the Metro 200 growing medium. Plants are watered daily before transplantation and approximately 3 times a week after transplantation. Peters 15-16-17 fertilizer is applied approximately three times per week after transplanting at a concentration of 150 ppm N. Two to three times during the life time of the plant from transplanting to flowering a total of approximately 900 mg Fe is added to each pot. Maize plants are grown in the green house in approximately 15hr day/9hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is at the 6 leaf development stage. The root system is cut from maize plant and washed with water to free it from the soil. The tissue is then immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON004 cDNA library is generated from maize (B73 x Mo17, Illinois Foundation Seeds, Champaign, Illinois U.S.A.) total leaf tissue at the V6 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after

transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is at the 6-leaf development stage. The older, more juvenile leaves, which are in a basal position, as well as the younger, more adult leaves, which are more apical are cut at the base of the leaves. The leaves are then pooled and immediately transferred to liquid nitrogen containers in which the pooled leaves are crushed. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON005 cDNA library is generated from maize (B73 x Mo17, Illinois Foundation Seeds, Champaign Illinois, U.S.A.) root tissue at the V6 development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the green house in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is at the 6-leaf development stage. The root system is cut from the mature maize plant and washed with water to free it from

the soil. The tissue is immediately frozen in liquid nitrogen and the harvested tissue is then stored at -80°C until RNA preparation.

The SATMON006 cDNA library is generated from maize (B73 x Mo17, Illinois Foundation Seeds, Champaign Illinois, U.S.A.) total leaf tissue at the V6 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is at the 6-leaf development stage. The older more juvenile leaves, which are in a basal position, as well as the younger more adult leaves, which are more apical are cut at the base of the leaves. The leaves are then pooled and immediately transferred to liquid nitrogen containers in which the pooled leaves are crushed. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON007 cDNA library is generated from the primary root tissue of 5 day old maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) seedlings. Seeds are planted on a moist filter paper on a covered tray that is kept in the dark until germination (one day). After germination, the trays, along with the moist paper, are moved to a greenhouse where the maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles for approximately 5 days. The

daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. The primary root tissue is collected when the seedlings are 5 days old. At this stage, the primary root (radicle) is pushed through the coleorhiza which itself is pushed through the seed coat. The primary root, which is about 2-3 cm long, is cut and immediately frozen in liquid nitrogen and then stored at -80°C until RNA preparation.

The SATMON008 cDNA library is generated from the primary shoot (coleoptile 2-3 cm) of maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) seedlings which are approximately 5 days old. Seeds are planted on a moist filter paper on a covered tray that is kept in the dark until germination (one day). Then the trays containing the seeds are moved to a greenhouse at 15hr daytime/9 hr nighttime cycles and grown until they are 5 days post germination. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Tissue is collected when the seedlings are 5 days old. At this stage, the primary shoot (coleoptile) is pushed through the seed coat and is about 2-3 cm long. The coleoptile is dissected away from the rest of the seedling, immediately frozen in liquid nitrogen and then stored at -80°C until RNA preparation.

The SATMON009 cDNA library is generated from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) leaves at the 8 leaf stage (V8 plant development stage). Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a

strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the green house in 15 hr day/9 hr night cycles. The daytime temperature is 80°F and the nighttime temperature is 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is at the 8-leaf development stage. The older more juvenile leaves, which are in a basal position, as well as the younger more adult leaves, which are more apical, are cut at the base of the leaves. The leaves are then pooled and then immediately transferred to liquid nitrogen containers in which the pooled leaves are crushed. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON010 cDNA library is generated from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) root tissue at the V8 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the green house in 15 hr day/9 hr night cycles. The daytime temperature is 80°F and the nighttime temperature is 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is at the V8 development stage. The root system is cut from this mature maize plant and washed with water to free it from the soil. The tissue is immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON011 cDNA library is generated from undeveloped maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) leaf at the V6 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the green house in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is at the 6-leaf development stage. The second youngest leaf which is at the base of the apical leaf of V6 stage maize plant is cut at the base and immediately transferred to liquid nitrogen containers in which the leaf is crushed. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON012 cDNA library is generated from 2 day post germination maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) seedlings. Seeds are planted on a moist filter paper on a covered tray that is kept in the dark until germination (one day). Then the trays containing the seeds are moved to the greenhouse and grown at 15hr daytime/9 hr nighttime cycles until 2 days post germination. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Tissue is collected when the seedlings are 2 days old. At the two day stage, the coleorhiza is pushed through the seed coat and the primary root (the radicle) is pierced the coleorhiza but is barely visible. Also, at this two day stage, the

coleoptile is just emerging from the seed coat. The 2 days post germination seedlings are then immersed in liquid nitrogen and crushed. The harvested tissue is stored at -80°C until preparation of total RNA.

The SATMON013 cDNA library is generated from apical maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) meristem founder at the V4 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Prior to tissue collection, the plant is at the 4 leaf stage. The lead at the apex of the V4 stage maize plant is referred to as the meristem founder. This apical meristem founder is cut, immediately frozen in liquid nitrogen and crushed. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON014 cDNA library is generated from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) endosperm fourteen days after pollination. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation.

Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. After the V10 stage, the maize plant ear shoots are ready for fertilization. At this stage, the ear shoots are enclosed in a paper bag before silk emergence to withhold the pollen. The ear shoots are pollinated and 14 days after pollination, the ears are pulled out and then the kernels are plucked out of the ears. Each kernel is then dissected into the embryo and the endosperm and the aleurone layer is removed. After dissection, the endosperms are immediately frozen in liquid nitrogen and then stored at -80°C until RNA preparation.

The SATMON016 library is a maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) sheath library collected at the V8 developmental stage. Seeds are planted in a depth of approximately 3 cm in solid into 2-3 inch pots containing Metro growing medium. After 2-3 weeks growth, they are transplanted into 10" pots containing the same. Plants are watered daily before transplantation and approximately the times a week after transplantation. Peters 15-16-17 fertilizer is applied approximately three times per week after transplanting, at a strength of 150 ppm N. Two to three times during the life time of the plant from transplanting to flowering, a total of approximately 900 mg Fe is added to each pot. Maize plants are grown in the green house in 15hr day/9hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. When the maize plants are at the V8 stage the 5<sup>th</sup> and 6<sup>th</sup> leaves from the



bottom exhibit fully developed leaf blades. At the base of these leaves, the ligule is differentiated and the leaf blade is joined to the sheath. The sheath is dissected away from the base of the leaf then the sheath is frozen in liquid nitrogen and crushed. The tissue is then stored at -80°C until RNA preparation.

The SATMON017 cDNA library is generated from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) embryo seventeen days after pollination. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth the seeds are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the green house in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. After the V10 stage, the ear shoots of maize plant, which are ready for fertilization, are enclosed in a paper bag before silk emergence to withhold the pollen. The ear shoots are fertilized and 21 days after pollination, the ears are pulled out and the kernels are plucked out of the ears. Each kernel is then dissected into the embryo and the endosperm and the aleurone layer is removed. After dissection, the embryos are immediately frozen in liquid nitrogen and then stored at -80°C until RNA preparation.

The SATMON019 (Lib3054) cDNA library is generated from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) culm (stem) at the V8 developmental stage. Seeds are planted

at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the green house in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. When the maize plant is at the V8 stage, the 5th and 6th leaves from the bottom have fully developed leaf blades. The region between the nodes of the 5th and the sixth leaves from the bottom is the region of the stem that is collected. The leaves are pulled out and the sheath is also torn away from the stem. This stem tissue is completely free of any leaf and sheath tissue. The stem tissue is then frozen in liquid nitrogen and stored at -80°C until RNA preparation.

The SATMON020 cDNA library is from a maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) Hill Type II-Initiated Callus. Petri plates containing approximately 25 ml of Type II initiation media are prepared. This medium contains N6 salts and vitamins, 3% sucrose, 2.3 g/liter proline 0.1 g/liter enzymatic casein hydrolysate, 2mg/liter 2,4 – dichloro phenoxy-acetic acid (2,4, D), 15.3 mg/liter AgNO<sub>3</sub> and 0.8% bacto agar and is adjusted to pH 6.0 before autoclaving. At 9-11 days after pollination, an ear with immature embryos measuring approximately 1-2 mm in length is chosen. The husks and silks are removed and then the ear is broken into halves and placed in an autoclaved solution of Clorox/TWEEN 20 sterilizing solution. Then the ear is rinsed with deionized water. Then each embryo is extracted from the

kernel. Intact embryos are placed in contact with the medium, scutellar side up). Multiple embryos are plated on each plate and the plates are incubated in the dark at 25°C. Type II calluses are friable, can be subcultured with a spatula, frequently regenerate via somatic embryogenesis and are relatively undifferentiated. As seen in the microscope, the Type II calluses show color ranging from translucent to light yellow and heterogeneity with respect to embryoid structure as well as stage of embryoid development. Once Type II callus are formed, the callus is transferred to type II callus maintenance medium without  $\text{AgNO}_3$ . Every 7-10 days, the callus is subcultured. About 4 weeks after embryo isolation the callus is removed from the plates and then frozen in liquid nitrogen. The harvested tissue is stored at  $-80^\circ\text{C}$  until RNA preparation.

The SATMON021 cDNA library is generated from the immature maize (DK604, Dekalb Genetics, Dekalb Illinois, U.S.A.) tassel at the V8 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the green house in 15 hr day/9 hr night cycles. The daytime temperature is approximately  $80^\circ\text{F}$  and the nighttime temperature is approximately  $70^\circ\text{F}$ . Supplemental lighting is provided by 1000 W sodium vapor lamps. As the maize plant enters the V8 stage, tassels which are 15-20 cm in length are collected and frozen in liquid nitrogen. The harvested tissue is stored at  $-80^\circ\text{C}$  until RNA preparation.

6034560  
The SATMON022 cDNA library is generated from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) ear (growing silks) at the V8 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. *Zea mays* plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the plant is in the V8 stage. At this stage, some immature ear shoots are visible. The immature ear shoots (approximately 1 cm in length) are pulled out, frozen in liquid nitrogen and then stored at -80°C until RNA preparation.

The SATMON23 cDNA library is generated from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) ear (growing silk) at the V8 development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime

temperature is approximately 70°F. When the tissue is harvested at the V8 stage, the length of the ear that is harvested is about 10-15 cm and the silks are just exposed (approximately 1 inch). The ear along with the silks is frozen in liquid nitrogen and then the tissue is stored at -80°C until RNA preparation.

The SATMON024 cDNA library is generated from the immature maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) tassel at the V9 development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the green house in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. As a maize plant enters the V9 stage, the tassel is rapidly developing and a 37 cm tassel along with the glume, anthers and pollen is collected and frozen in liquid nitrogen. The harvested tissue is stored at -80°C until RNA preparation.

The SATMON025 cDNA library is from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) Hill Type II-Regenerated Callus. Type II callus is grown in initiation media as described for SATMON020 and then the embryoids on the surface of the Type II callus are allowed to mature and germinate. The 1-2 gm fresh weight of the soft friable type callus containing numerous embryoids are transferred to 100 x 15 mm petri plates containing 25 ml of regeneration media. Regeneration media consists of Murashige and Skoog (MS) basal salts,

modified White's vitamins (0.2 g/liter glycine and 0.5 g/liter myo-inositol and 0.8% bacto agar (6SMS0D)). The plates are then placed in the dark after covering with parafilm. After 1 week, the plates are moved to a lighted growth chamber with 16 hr light and 8 hr dark photoperiod. Three weeks after plating the Type II callus to 6SMS0D, the callus exhibit shoot formation. The callus and the shoots are transferred to fresh 6SMS0D plates for another 2 weeks. The callus and the shoots are then transferred to petri plates with reduced sucrose (3SMS0D). Upon distinct formation of a root and shoot, the newly developed green plants are then removed out with a spatula and frozen in liquid nitrogen containers. The harvested tissue is then stored at  $-80^{\circ}\text{C}$  until RNA preparation.

The SATMON026 cDNA library is generated from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) juvenile/adult shift leaves at the V8 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the green house in 15 hr day/9 hr night cycles. The daytime temperature is approximately  $80^{\circ}\text{F}$  and the nighttime temperature is approximately  $70^{\circ}\text{F}$ . Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plants are at the 8-leaf development stage. Leaves are founded sequentially around the meristem over weeks of time and the older, more juvenile leaves arise earlier and in a more basal position than the younger, more adult leaves, which are in a

more apical position. In a V8 plant, some leaves which are in the middle portion of the plant exhibit characteristics of both juvenile as well as adult leaves. They exhibit a yellowing color but also exhibit, in part, a green color. These leaves are termed juvenile/adult shift leaves. The juvenile/adult shift leaves (the 4th, 5th leaves from the bottom) are cut at the base, pooled and transferred to liquid nitrogen in which they are then crushed. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON027 cDNA library is generated from 6 day maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) leaves. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the Metro 200 growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. *Zea mays* plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Prior to tissue collection, when the plant is at the 8-leaf stage, water is held back for six days. The older, more juvenile leaves, which are in a basal position, as well as the younger, more adult leaves, which are more apical, are all cut at the base of the leaves. All the leaves exhibit significant wilting. The leaves are then pooled and immediately transferred to liquid nitrogen containers in which the pooled leaves are then crushed. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON028 cDNA library is generated from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) roots at the V8 developmental stage that are subject to six days water stress. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the Metro 200 growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Prior to tissue collection, when the plant is at the 8-leaf stage, water is held back for six days. The root system is cut, shaken and washed to remove soil. Root tissue is then pooled and immediately transferred to liquid nitrogen containers in which the pooled leaves are then crushed. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON029 cDNA library is generated from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) seedlings at the etiolated stage. Seeds are planted on a moist filter paper on a covered tray that is kept in the dark for 4 days at approximately 70°F. Tissue is collected when the seedlings are 4 days old. By 4 days, the primary root has penetrated the coleorhiza and is about 4-5 cm and the secondary lateral roots have also made their appearance. The coleoptile has also pushed through the seed coat and is about 4-5 cm long. The seedlings are frozen in liquid nitrogen and crushed. The harvested tissue is then stored at -80°C until RNA preparation.



bioRxiv preprint doi: <https://doi.org/10.1101/007260>; this version posted February 2, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The SATMON030 cDNA library is generated from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) root tissue at the V4 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth, they are transplanted into 10 inch pots containing the same. Plants are watered daily before transplantation and approximately 3 times a week after transplantation. Peters 15-16-17 fertilizer is applied approximately three times per week after transplanting, at a strength of 150 ppm N. Two to three times during the life time of the plant, from transplanting to flowering, a total of approximately 900 mg Fe is added to each pot. Maize plants are grown in the green house in 15hr day/9hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 sodium vapor lamps. Tissue is collected when the maize plant is at the 4 leaf development stage. The root system is cut from the mature maize plant and washed with water to free it from the soil. The tissue is then immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON031 cDNA library is generated from the maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) leaf tissue at the V4 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the green house

in 15 hr day/9 hr night cycles. The daytime temperature is 80°F and the nighttime temperature is 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is at the 4-leaf development stage. The third leaf from the bottom is cut at the base and immediately frozen in liquid nitrogen and crushed. The tissue is immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON033 cDNA library is generated from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) embryo tissue 13 days after pollination. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. After the V10 stage, the ear shoots of the maize plant, which are ready for fertilization, are enclosed in a paper bag before silk emergent to withhold the pollen. The ear shoots are pollinated and 13 days after pollination, the ears are pulled out and then the kernels are plucked cut of the ears. Each kernel is then dissected into the embryo and the endosperm and the aleurone layer is removed. After dissection, the embryos are immediately frozen in liquid nitrogen and then stored at -80°C until RNA preparation.

The SATMON034 cDNA library is generated from cold stressed maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) seedlings. Seeds are planted on a moist filter paper on a covered tray that is kept on at 10°C for 7 days. After 7 days, the temperature is shifted to 15°C for one day until germination of the seed. Tissue is collected once the seedlings are 1 day old. At this point, the coleorhiza has just pushed out of the seed coat and the primary root is just making its appearance. The coleoptile has not yet pushed completely through the seed coat and is also just making its appearance. These 1 day old cold stressed seedlings are frozen in liquid nitrogen and crushed. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMON~001 (Lib36, Lib83, Lib84) cDNA library is generated from maize leaves at the V8 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in a greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue from the maize plant is collected at the V8 stage. The older more juvenile leaves in a basal position as well as the younger more adult leaves which are more apical are all cut at the base, pooled and frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SATMONN01 cDNA library is generated from maize (B73, Illinois Foundation Seeds, Champaign, Illinois U.S.A.) normalized immature tassels at the V6 plant development stage normalized tissue. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in a greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue from the maize plant is collected at the V6 stage. At that stage the tassel is an immature tassel of about 2-3 cm in length. The tassels are removed and frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation. Single stranded and double stranded DNA representing approximately  $1 \times 10^6$  colony forming units are isolated using standard protocols. RNA, complementary to the single stranded DNA, is synthesized using the double stranded DNA as a template. Biotinylated dATP is incorporated into the RNA during the synthesis reaction. The single stranded DNA is mixed with the biotinylated RNA in a 1:10 molar ratio) and allowed to hybridize. DNA-RNA hybrids are captured on Dynabeads M280 streptavidin (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The dynabeads with captured hybrids are collected with a magnet. The non-hybridized single stranded molecules remaining after hybrid capture are converted to double stranded form and represent the primary normalized library.

The SATMONN04 cDNA library is generated from maize (B73 x Mo17, Illinois Foundation Seeds, Champaign, Illinois U.S.A.) normalized total leaf tissue at the V6 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is at the 6-leaf development stage. The older, more juvenile leaves, which are in a basal position, as well as the younger, more adult leaves, which are more apical are cut at the base of the leaves. The leaves are then pooled and immediately transferred to liquid nitrogen containers in which the pooled leaves are crushed. The harvested tissue is then stored at -80°C until RNA preparation. Single stranded and double stranded DNA representing approximately  $1 \times 10^6$  colony forming units are isolated using standard protocols. RNA, complementary to the single stranded DNA, is synthesized using the double stranded DNA as a template. Biotinylated dATP is incorporated into the RNA during the synthesis reaction. The single stranded DNA is mixed with the biotinylated RNA in a 1:10 molar ratio) and allowed to hybridize. DNA-RNA hybrids are captured on Dynabeads M280 streptavidin (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The dynabeads with captured hybrids are collected with a magnet. The non-

hybridized single stranded molecules remaining after hybrid capture are converted to double stranded form and represent the primary normalized library.

The SATMONN05 cDNA library is generated from maize (B73 x Mo17, Illinois Foundation Seeds, Champaign Illinois, U.S.A.) normalized root tissue at the V6 development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the green house in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is at the 6-leaf development stage. The root system is cut from the mature maize plant and washed with water to free it from the soil. The tissue is immediately frozen in liquid nitrogen and the harvested tissue is then stored at -80°C until RNA preparation. The single stranded and double stranded DNA representing approximately  $1 \times 10^6$  colony forming units are isolated using standard protocols. RNA, complementary to the single stranded DNA, is synthesized using the double stranded DNA as a template. Biotinylated dATP is incorporated into the RNA during the synthesis reaction. The single stranded DNA is mixed with the biotinylated RNA in a 1:10 molar ratio) and allowed to hybridize. DNA-RNA hybrids are captured on Dynabeads M280 streptavidin (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The dynabeads

with captured hybrids are collected with a magnet. The non-hybridized single stranded molecules remaining after hybrid capture are converted to double stranded form and represent the primary normalized library.

The SATMONN06 cDNA library is generated from maize (B73 x Mo17, Illinois Foundation Seeds, Champaign Illinois, U.S.A.) normalized total leaf tissue at the V6 plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is at the 6-leaf development stage. The older more juvenile leaves, which are in a basal position, as well as the younger more adult leaves, which are more apical are cut at the base of the leaves. The leaves are then pooled and immediately transferred to liquid nitrogen containers in which the pooled leaves are crushed. The harvested tissue is then stored at -80°C until RNA preparation. Single stranded and double stranded DNA representing approximately  $1 \times 10^6$  colony forming units are isolated using standard protocols. RNA, complementary to the single stranded DNA, is synthesized using the double stranded DNA as a template. Biotinylated dATP is incorporated into the RNA during the synthesis reaction. The single stranded DNA is mixed with the

biotinylated RNA in a 1:10 molar ratio) and allowed to hybridize. DNA-RNA hybrids are captured on Dynabeads M280 streptavidin (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The dynabeads with captured hybrids are collected with a magnet. The non-hybridized single stranded molecules remaining after hybrid capture are converted to double stranded form and represent the primary normalized library.

The CMZ029 (SATMON036) cDNA library is generated from maize (DK604, Dekalb Genetics, Dekalb, Illinois U.S.A.) endosperm 22 days after pollination. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the green house in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. After the V10 stage, the ear shoots of the maize plant, which are ready for fertilization, are enclosed in a paper bag before silk emergent to withhold the pollen. The ear shoots are pollinated and 22 days after pollination, the ears are pulled out and then the kernels are plucked out of the ears. Each kernel is then dissected into the embryo and the endosperm and the alurone layer is removed. After dissection, the endosperms are immediately frozen in liquid nitrogen and then stored at -80°C until RNA preparation.



The CMz030 (Lib143) cDNA library is generated from maize seedling tissue two days post germination. Seeds are planted on a moist filter paper on a covered tray that is kept in the dark until germination. The trays are then moved to the bench top at 15 hr daytime/9 hr nighttime cycles for 2 days post-germination. The day time temperature is 80°F and the nighttime temperature is 70°F. Tissue is collected when the seedlings are 2 days old. At this stage, the colehrhiza has pushed through the seed coat and the primary root (the radicle) is just piercing the colehrhiza and is barely visible. The seedlings are placed at 42°C for 1 hour. Following the heat shock treatment, the seedlings are immersed in liquid nitrogen and crushed. The harvested tissue is stored at -80° until RNA preparation.

The CMz031 (Lib148) cDNA library is generated from maize pollen tissue at the V10+ plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected from V10+ stage plants. The ear shoots, which are ready for fertilization, are enclosed in a paper bag to withhold pollen. Twenty-one days after pollination, prior to removing the ears, the paper bag is shaken to collect the mature pollen. The mature pollen is immediately frozen in liquid

nitrogen containers and the pollen is crushed. The harvested tissue is then stored at -80°C until RNA preparation.

The CMz033 (Lib189) cDNA library is generated from maize pooled leaf tissue. Samples are harvested from open pollinated plants. Tissue is collected from maize leaves at the anthesis stage. The leaves are collected from 10-12 plants and frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The CMz034 (Lib3060) cDNA library is generated from maize mature tissue at 40 days post pollination plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected from leaves located two leaves below the ear leaf. This sample represents those genes expressed during onset and early stages of leaf senescence. The leaves are pooled and immediately transferred to liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The CMz035 (Lib3061) cDNA library is generated from maize endosperm tissue at the V10+ plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch

peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80 F and the nighttime temperature is approximately 70 F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected from V10+ stage plants. The ear shoots, which are ready for fertilization, are enclosed in a paper bag prior to silk emergence to withhold pollen. Thirty-two days after pollination, the ears are pulled out and the kernels are removed from the cob. Each kernel is dissected into the embryo and the endosperm and the aleurone layer is removed. After dissection, the endosperms are immediately transferred to liquid nitrogen. The harvested tissue is then stored at -80 C until RNA preparation.

The CMz036 (Lib3062) cDNA library is generated from maize husk tissue at the 8 week old plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F.

Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected from 8 week old plants. The husk is separated from the ear and immediately transferred to liquid nitrogen containers. The harvested tissue is then stored at -80°C until RNA preparation.

The CMz037 (Lib3059) cDNA library is generated from maize pooled kernal at 12-15 days after pollination plant development stage. Sample were collected from field grown material. Whole kernal from hand pollinated (control pollination) are harvested as whole ears and immediately frozen on dry ice. Kernels from 10-12 ears were pooled and ground together in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The CMz039 (Lib3066) cDNA library is generated from maize immature anther tissue at the 7 week old immature tassel stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F.

Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is at the 7 week old immature tassel stage. At this stage, prior to anthesis, the immature anthers are green and enclosed in the staminate spikelet. The developing anthers are dissected away from the 7 week old immature tassel and immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The CMz040 (Lib3067) cDNA library is generated from maize kernel tissue at the V10+ plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected from V10+ stage plants. The ear shoots, which are ready for fertilization, are enclosed in a paper bag before silk emergence to withhold pollen. Five to eight days after controlled pollination. The ears are pulled and the kernels removed. The kernels are immediately frozen in liquid nitrogen. The harvested kernels tissue is then stored at -80°C until RNA preparation. This sample represents gene expressed in early kernel development, during periods of cell division, amyloplast biogenesis and early carbon flow across the material to filial tissue.

The CMz041 (Lib3068) cDNA library is generated from maize pollen germinating silk tissue at the V10+ plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times

during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F.

Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected from V10+ stage plants when the ear shoots are ready for fertilization at the silk emergence stage. The emerging silks are pollinated with an excess of pollen under controlled pollination conditions in the green house. Eighteen hours after pollination the silks are removed from the ears and immediately frozen in liquid nitrogen containers. This sample represents genes expressed in both pollen and silk tissue early in pollination. The harvested tissue is then stored at -80°C until RNA preparation.

The CMz042 (Lib3069) cDNA library is generated from maize ear tissue excessively pollinated at the V10+ plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F.

Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected from V10+ stage plants and the ear shoots which are ready for fertilization are at the silk emergence stage. The immature ears are pollinated with an excess of pollen under controlled pollination

conditions. Eighteen hours post-pollination, the ears are removed and immediately transferred to liquid nitrogen containers. The harvested tissue is then stored at -80°C until RNA preparation.

The CMz044 (Lib3075) cDNA library is generated from maize microspore tissue at the V10+ plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected from immature anthers from 7 week old tassels. The immature anthers are first dissected from the 7 week old tassel with a scalpel on a glass slide covered with water. The microspores (immature pollen) are released into the water and are recovered by centrifugation. The microspore suspension is immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The CMz045 (Lib3076) cDNA library is generated from maize immature ear megaspore tissue. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after

transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected from immature ear (megaspore) obtained from 7 week old plants. The immature ears are harvested from the 7 week old plants and are approximately 2.5 to 3 cm in length. The kernels are removed from the cob immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The CMz047 (Lib3078) cDNA library is generated from maize CO<sub>2</sub> treated high-exposure shoot tissue at the V10+ plant development stage. RX601 maize seeds are sterilized for 1 minute with a 10% clorox solution. The seeds are rolled in germination paper, and germinated in 0.5 mM calcium sulfate solution for two days at 30°C. The seedlings are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium at a rate of 2-3 seedlings per pot. Twenty pots are placed into a high CO<sub>2</sub> environment (approximately 1000 ppm CO<sub>2</sub>). Twenty plants were grown under ambient greenhouse CO<sub>2</sub> (approximately 450 ppm CO<sub>2</sub>). Plants are watered daily before transplantation and three times a week after transplantation. Peters 20-20-20 fertilizer is also lightly applied. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. At ten days post planting, the shoots from both atmosphere are frozen in liquid nitrogen and lightly ground. The roots are washed in deionized water to remove the support media and



the tissue is immediately transferred to liquid nitrogen containers. The harvested tissue is then stored at -80°C until RNA preparation.

The CMz048 (Lib3079) cDNA library is generated from maize basal endosperm transfer layer tissue at the V10+ plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected from V10+ maize plants. The ear shoots, which are ready for fertilization, are enclosed in a paper bag prior to silk emergence, to withhold the pollen. Kernels are harvested at 12 days post-pollination and placed on wet ice for dissection. The kernels are cross sectioned laterally, dissecting just above the pedicel region, including 1-2 mm of the lower endosperm and the basal endosperm transfer region. The pedicel and lower endosperm region containing the basal endosperm transfer layer is pooled and immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The CMz049(Lib3088) cDNA library is generated from maize immature anther tissue at the 7 week old immature tassel stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are

transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is at the 7 week old immature tassel stage. At this stage, prior to anthesis, the immature anthers are green and enclosed in the staminate spikelet. The developing anthers are dissected away from the 7 week old immature tassel and immediately transferred to liquid nitrogen container. The harvested tissue is then stored at -80°C until RNA preparation.

The CMz050 (Lib3114) cDNA library is generated from maize silk tissue at the V10+ plant development stage. Seeds are planted at a depth of approximately 3 cm into 2-3 inch peat pots containing Metro 200 growing medium. After 2-3 weeks growth they are transplanted into 10 inch pots containing the same growing medium. Plants are watered daily before transplantation and three times a week after transplantation. Peters 15-16-17 fertilizer is applied three times per week after transplanting at a strength of 150 ppm N. Two to three times during the lifetime of the plant, from transplanting to flowering, a total of 900 mg Fe is added to each pot. Maize plants are grown in the greenhouse in 15 hr day/9 hr night cycles. The daytime temperature is approximately 80°F and the nighttime temperature is approximately 70°F. Supplemental lighting is provided by 1000 W sodium vapor lamps. Tissue is collected when the maize plant is beyond the 10-leaf development stage and the ear shoots are approximately 15-20

cm in length. The ears are pulled and silks are separated from the ears and immediately transferred to liquid nitrogen containers. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON001 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) total leaf tissue at the V4 plant development stage. Leaf tissue from 38, field grown V4 stage plants is harvested from the 4<sup>th</sup> node. Leaf tissue is removed from the plants and immediately frozen in dry-ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON002 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) root tissue at the V4 plant development stage. Root tissue from 76, field grown V4 stage plants is harvested. The root systems is cut from the soybean plant and washed with water to free it from the soil and immediately frozen in dry-ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON003 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) seedling hypocotyl axis tissue harvested 2 day post-imbibition. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium. Trays are placed in an environmental chamber and grown at 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature approximately 24°C. Soil is checked and watered daily to maintain even moisture conditions. Tissue is collected 2 days after the start of imbibition. The 2 days after imbibition samples are separated into 3 collections after removal of any adhering seed coat. At the 2 day stage, the hypocotyl axis is emerging from the soil. A few seedlings have cracked the

soil surface and exhibited slight greening of the exposed cotyledons. The seedlings are washed in water to remove soil, hypocotyl axis harvested and immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON004 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) seedling cotyledon tissue harvested 2 day post-imbibition. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium. Trays are placed in an environmental chamber and grown at 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature approximately 24°C. Soil is checked and watered daily to maintain even moisture conditions. Tissue is collected 2 days after the start of imbibition. The 2 days after imbibition samples are separated into 3 collections after removal of any adhering seed coat. At the 2 day stage, the hypocotyl axis is emerging from the soil. A few seedlings have cracked the soil surface and exhibited slight greening of the exposed cotyledons. The seedlings are washed in water to remove soil, hypocotyl axis harvested and immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON005 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) seedling hypocotyl axis tissue harvested 6 hour post-imbibition. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium. Trays are placed in an environmental chamber and grown at 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature approximately 24°C. Soil is checked and watered daily to maintain even moisture conditions. Tissue is collected 6 hours after the start of imbibition. The 6 hours after imbibition samples are separated into 3 collections after removal of any adhering seed coat. The

6 hours after imbibition sample is collected over the course of approximately 2 hours starting at 6 hours post imbibition. At the 6 hours after imbibition stage, not all cotyledons have become fully hydrated and germination, or radicle protrusion, has not occurred. The seedlings are washed in water to remove soil, hypocotyl axis harvested and immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON006 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) seedling cotyledons tissue harvest 6 hour post-imbibition. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium. Trays are placed in an environmental chamber and grown at 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature approximately 24°C. Soil is checked and watered daily to maintain even moisture conditions. Tissue is collected 6 hours after imbibition. The 6 hours after imbibition samples are separated into 3 collections after removal of any adhering seed coat. The 6 hours after imbibition sample is collected over the course of approximately 2 hours starting at 6 hours post-imbibition. At the 6 hours after imbibition, not all cotyledons have become fully hydrated and germination or radicle protrusion, have not occurred. The seedlings are washed in water to remove soil, cotyledon harvested and immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON007 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) seed tissue harvested 25 and 35 days post-flowering. Seed pods from field grown plants are harvested 25 and 35 days after flowering and the seeds extracted from the pods. Approximately 4.4g and 19.3g of seeds are harvested from the

respective seed pods and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON008 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) leaf tissue harvested from 25 and 35 days post-flowering plants. Total leaf tissue is harvested from field grown plants. Approximately 19g and 29g of leaves are harvested from the fourth node of the plant 25 and 35 days post-flowering and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON009 cDNA library is generated from soybean cutlivar C1944 (USDA Soybean Germplasm Collection, Urbana, Illinois U.S.A.) pod and seed tissue harvested 15 days post-flowering. Pods from field grown plants are harvested 15 days post-flowering. Approximately 3g of pod tissue is harvested and immediately frozen in dry-ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON010 cDNA library is generated from soybean cultivar C1944 (USDA Soybean Germplasm Collection, Urbana, Illinois U.S.A.) seed tissue harvested 40 days post-flowering. Pods from field grown plants are harvested 40 days post-flowering. Pods and seeds are separated, approximately 19g of seed tissue is harvested and immediately frozen in dry-ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON011 cDNA library is generated from soybean cultivars Cristalina (USDA Soybean Germplasm Collection, Urbana, Illinois U.S.A.) and FT108 (Monsoy, Brazil) (tropical germ plasma) leaf tissue. Leaves are harvested from plants grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and

the nighttime temperature approximately 24°C. Soil is checked and watered daily to maintain even moisture conditions. Approximately 30g of leaves are harvested from the 4<sup>th</sup> node of each of the Cristalina and FT108 cultivars and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON012 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) leaf tissue. Leaves from field grown plants are harvested from the fourth node 15 days post-flowering. Approximately 12g of leaves are harvested and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON013 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) root and nodule tissue. Approximately, 28g of root tissue from field grown plants is harvested 15 days post-flowering. The root system is cut from the soybean plant, washed with water to free it from the soil and immediately frozen in dry-ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON014 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) seed tissue harvested 25 and 35 days after flowering. Seed pods from field grown plants are harvested 15 days after flowering and the seeds extracted from the pods. Approximately 5g of seeds are harvested from the respective seed pods and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON015 cDNA is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) seed tissue harvested 45 and 55 days post-flowering. Seed pods from field grown plants are harvested 45 and 55 days after flowering and the seeds

extracted from the pods. Approximately 19g and 31g of seeds are harvested from the respective seed pods and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON016 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) root tissue. Approximately, 61g and 38g of root tissue from field grown plants is harvested 25 and 35 days post- flowering is harvested. The root system is cut from the soybean plant and washed with water to free it from the soil. The tissue is placed in 14ml polystyrene tubes and immediately frozen in dry-ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON017 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) root tissue. Approximately 28g of root tissue from field grown plants is harvested 45 and 55 days post-flowering. The root system is cut from the soybean plant, washed with water to free it from the soil and immediately frozen in dry-ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON018 cDNA is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) leaf tissue harvested 45 and 55 days post-flowering. Leaves from field grown plants are harvested 45 and 55 days after flowering from the fourth node. Approximately 27g and 33g of seeds are harvested from the respective seed pods and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON019 cDNA library is generated from soybean cultivars Cristalina (USDA Soybean Germplasm Collection, Urbana, Illinois U.S.A.) and FT108 (Monsoy, Brazil) (tropical germ plasma) root tissue. Roots are harvested from plants grown in an environmental chamber



under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature approximately 24°C. Soil is checked and watered daily to maintain even moisture conditions. Approximately 50g and 56g of roots are harvested from each of the Cristalina and FT108 cultivars and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON020 cDNA is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) seed tissue harvested 65 and 75 days post-flowering. Seed pods from field grown plants are harvested 45 and 55 days after flowering and the seeds extracted from the pods. Approximately 14g and 31g of seeds are harvested from the respective seed pods and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON021 cDNA library is generated from Soybean Cyst Nematode-resistant soybean cultivar Hartwig (USDA Soybean Germplasm Collection, Urbana, Illinois U.S.A.) root tissue. Plants are grown in tissue culture at room temperature. At approximately 6 weeks post-germination, the plants are exposed to sterilized Soybean Cyst Nematode eggs. Infection is then allowed to progress for 10 days. After the 10 day infection process, the tissue is harvested. Agar from the culture medium and nematodes are removed and the root tissue is immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON022 (Lib3030) cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) partially opened flower tissue. Partially to fully opened flower tissue is harvested from plants grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature approximately 24°C. Soil is checked and watered daily to

maintain even moisture conditions. A total of 3g of flower tissue is harvested and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON023 cDNA library is generated from soybean genotype BW211S Null (Tohoku University, Morioka, Japan) seed tissue harvested 15 and 40 days post-flowering. Seed pods from field grown plants are harvested 15 and 40 days post-flowering and the seeds extracted from the pods. Approximately 0.7g and 14.2g of seeds are harvested from the respective seed pods and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON024 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) internode-2 tissue harvested 18 days post-imbibition. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium. The plants are grown in a greenhouse for 18 days after the start of imbibition at ambient temperature. Soil is checked and watered daily to maintain even moisture conditions. Stem tissue is harvested 18 days after the start of imbibition. The samples are divided into hypocotyl and internodes 1 through 5. The fifth internode contains some leaf bud material. Approximately 3 g of each sample is harvested and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON025 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) leaf tissue harvested 65 days post-flowering. Leaves are harvested from the fourth node of field grown plants 65 days post-flowering. Approximately 18.4g of leaf tissue is harvested and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

SOYMON026 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) root tissue harvested 65 and 75 days post-flowering. Approximately 27g and 40g of root tissue from field grown plants is harvested 65 and 75 days post- flowering. The root system is cut from the soybean plant, washed with water to free it from the soil and immediately frozen in dry-ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON027 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) seed tissue harvested 25 days post-flowering. Seed pods from field grown plants are harvested 25 days post-flowering and the seeds extracted from the pods. Approximately 17g of seeds are harvested from the seed pods and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON028 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) drought-stressed root tissue. The plants are grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature 24°C. Soil is checked and watered daily to maintain even moisture conditions. At the R3 stage of development, water is withheld from half of the plant collection (drought stressed population). After 3 days, half of the plants from the drought stressed condition and half of the plants from the control population are harvested. After another 3 days (6 days post drought induction) the remaining plants are harvested. A total of 27g and 40g of root tissue is harvested and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON029 cDNA library is generated from Soybean Cyst Nematode-resistant soybean cultivar PI07354 (USDA Soybean Germplasm Collection, Urbana, Illinois U.S.A.) root tissue. Late fall to early winter greenhouse grown plants are exposed to Soybean Cyst Nematode eggs. At 10 days post-infection, the plants are uprooted, rinsed briefly and the roots frozen in liquid nitrogen. Approximately 20 grams of root tissue is harvested from the infected plants. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON030 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) flower bud tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature approximately 24°C. Soil is checked and watered daily to maintain even moisture conditions. Flower buds are removed from the plant at the pedicel. A total of 100mg of flower buds are harvested and immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON031 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) carpel and stamen tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature approximately 24°C. Soil is checked and watered daily to maintain even moisture conditions. Flower buds are removed from the plant at the pedicel. Flowers are dissected to separate petals, sepals and reproductive structures (carpels and stamens). A total of 300mg of carpel and stamen tissue are

harvested and immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON032 cDNA library is prepared from the Asgrow cultivar A4922 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) rehydrated dry soybean seed meristem tissue. Surface sterilized seeds are germinated in liquid media for 24 hours. The seed axis is then excised from the barely germinating seed, placed on tissue culture media and incubated overnight at 20°C in the dark. The supportive tissue is removed from the explant prior to harvest. Approximately 570mg of tissue is harvested and frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON033 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) heat-shocked seedling tissue without cotyledons. Seeds are imbibed and germinated in vermiculite for 2 days under constant illumination. After 48 hours, the seedlings are transferred to an incubator set at 40°C under constant illumination. After 30, 60 and 180 minutes seedlings are harvested and dissected. A portion of the seedling consisting of the root, hypocotyl and apical hook is frozen in liquid nitrogen and stored at -80°C. The seedlings after 2 days of imbibition are beginning to emerge from the vermiculite surface. The apical hooks are dark green in appearance. Total RNA and poly A<sup>+</sup> RNA is prepared from equal amounts of pooled tissue.

The SOYMON034 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) cold-shocked seedling tissue without cotyledons. Seeds are imbibed and germinated in vermiculite for 2 days under constant illumination. After 48 hours, the seedlings are transferred to a cold room set at 5°C under constant illumination. After 30, 60 and 180 minutes seedlings are harvested and dissected. A

portion of the seedling consisting of the root, hypocotyl and apical hook is frozen in liquid nitrogen and stored at -80°C. The seedlings after 2 days of imbibition are beginning to emerge from the vermiculite surface. The apical hooks are dark green in appearance.

The SOYMON035 cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) seed coat tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature 24°C. Soil is checked and watered daily to maintain even moisture conditions. Seeds are harvested from mid to nearly full maturation (seed coats are not yellowing). The entire embryo proper is removed from the seed coat sample and the seed coat tissue are harvested and immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON036 cDNA library is generated from soybean cultivars PI171451, PI227687 and PI229358 (USDA Soybean Germplasm Collection, Urbana, Illinois U.S.A.) insect challenged leaves. Plants from each of the three cultivars are grown in screenhouse conditions. The screenhouse is divided in half and one half of the screenhouse is infested with soybean looper and the other half infested with velvetbean caterpillar. A single leaf is taken from each of the representative plants at 3 different time points, 11 days after infestation, 2 weeks after infestation and 5 weeks after infestation and immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation. Total RNA and poly A<sup>+</sup> RNA is isolated from pooled tissue consisting of equal quantities of all 18 samples (3 genotypes X 3 sample times X 2 insect genotypes).

The SOYMON037 cDNA library is generated from soybean cultivar A3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) etiolated axis and radical tissue. Seeds are planted in moist vermiculite, wrapped and kept at room temperature in complete darkness until harvest. Etiolated axis and hypocotyl tissue is harvested at 2, 3 and 4 days post-planting. A total of 1 gram of each tissue type is harvested at 2, 3 and 4 days after planting and immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The SOYMON038 cDNA library is generated from soybean variety Asgrow A3237 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) rehydrated dry seeds. Explants are prepared for transformation after germination of surface-sterilized seeds on solid tissue media. After 6 days, at 28°C and 18 hours of light per day, the germinated seeds are cold shocked at 4°C for 24 hours. Meristemic tissue and part of the hypocotyl is removed and cotyledon excised. The prepared explant is then wounded for *Agrobacterium* infection. The 2 grams of harvested tissue is frozen in liquid nitrogen and stored at -80°C until RNA preparation.

The Soy51 (LIB3027) cDNA library is prepared from equal amounts tissue harvested from SOYMON007, SOYMON015 and SOYMON020 prepared tissue. Single stranded and double stranded DNA representing approximately  $1 \times 10^6$  colony forming units are isolated using standard protocols. RNA, complementary to the single stranded DNA, is synthesized using the double stranded DNA as a template. Biotinylated dATP is incorporated into the RNA during the synthesis reaction. The single stranded DNA is mixed with the biotinylated RNA in a 1:10 molar ratio) and allowed to hybridize. DNA-RNA hybrids are captured on Dynabeads M280 streptavidin (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The dynabeads with captured hybrids are collected with a magnet. The non-hybridized single

stranded molecules remaining after hybrid capture are converted to double stranded form and represent the primary normalized library.

The Soy52 (LIB3028) cDNA library is generated from normalized flower DNA. Single stranded DNA representing approximately  $1 \times 10^6$  colony forming units of SOYMON022 harvested tissue is used as the starting material for normalization. RNA, complementary to the single stranded DNA, is synthesized using the double stranded DNA as a template. Biotinylated dATP is incorporated into the RNA during the synthesis reaction. The single stranded DNA is mixed with the biotinylated RNA in a 1:10 molar ratio) and allowed to hybridize. DNA-RNA hybrids are captured on Dynabeads M280 streptavidin (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The dynabeads with captured hybrids are collected with a magnet. The non-hybridized single stranded molecules remaining after hybrid capture are converted to double stranded form and represent the primary normalized library.

The Soy53 (LIB3039) cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) seedling shoot apical meristem tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature 24°C. Soil is checked and watered daily to maintain even moisture conditions. Apical tissue is harvested from seedling shoot meristem tissue, 7-8 days after the start of imbibition. The apex of each seedling is dissected to include the fifth node to the apical meristem. The fifth node corresponds to the third trifoliate leaf in the very early stages of development. Stipules completely envelop the leaf primordia at this time. A total of 200mg of apical tissue is harvested



and immediately frozen in liquid nitrogen. The harvested tissue is then stored at -80°C until RNA preparation.

The Soy54 (LIB3040) cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) heart to torpedo stage embryo tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature 24°C. Soil is checked and watered daily to maintain even moisture conditions. Seeds are collected and embryos removed from surrounding endosperm and maternal tissues. Embryos from globular to young torpedo stages (by corresponding analogy to *Arabidopsis*) are collected with a bias towards the middle of this spectrum. Embryos which are beginning to show asymmetric development of cotyledons are considered the upper developmental boundary for the collection and are excluded. A total of 12 mg embryo tissue is frozen in liquid nitrogen. The harvested tissue is stored at -80°C until RNA preparation.

Soy55 (LIB3049) cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) young seed tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature 24°C. Soil is checked and watered daily to maintain even moisture conditions. Seeds are collected from very young pods (5 to 15 days after flowering). A total of 100mg of seeds are harvested and frozen in liquid nitrogen. The harvested tissue is stored at -80°C until RNA preparation.

Soy56 (LIB3029) cDNA library is prepared from equal amounts tissue harvested from SOYMON007, SOYMON015 and SOYMON020 prepared tissue. Single stranded and double stranded DNA representing approximately  $1 \times 10^6$  colony forming units are isolated using standard protocols. RNA, complementary to the single stranded DNA, is synthesized using the double stranded DNA as a template. Biotinylated dATP is incorporated into the RNA during the synthesis reaction. The single stranded DNA is mixed with the biotinylated RNA in a 1:10 molar ratio and allowed to hybridize. DNA-RNA hybrids are captured on Dynabeads M280 streptavidin (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The dynabeads with captured hybrids are collected with a magnet. The non-hybridized single stranded molecules remaining after hybrid capture are not converted to double stranded form and represent a non-normalized seed pool for comparison to Soy51 cDNA libraries.

The Soy58 (LIB3050) cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) drought stressed root tissue subtracted from control root tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature 24°C. Soil is checked and watered daily to maintain even moisture conditions. At the R3 stage of the plant drought is induced by withholding water. After 3 and 6 days root tissue from both drought stressed and control (watered regularly) plants are collected and frozen in dry-ice. The harvested tissue is stored at -80°C until RNA preparation. For subtraction, target cDNA is made from the drought stressed tissue total RNA using the SMART cDNA synthesis system from Clontech (Clontech Laboratories, Palo Alto, California U.S.A.). Driver first strand cDNA is covalently linked to Dynabeads following a protocol similar to that

described in the Dynal literature (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The target cDNA is then heat denatured and the second strand trapped using Dynabeads oligo-dT. The target second strand cDNA is then hybridized to the driver cDNA in 400 l 2X SSPE for two rounds of hybridization at 65°C and 20 hours. After each hybridization, the hybridization solution is removed from the system and the hybridized target cDNA removed from the driver by heat denaturation in water. After hybridization, the remaining cDNA is trapped with Dynabeads oligo-dT. The trapped cDNA is then amplified as in previous PCR based libraries and the resulting cDNA ligated into the pSPORT vector (Invitrogen, Carlsbad California U.S.A.).

The Soy59 (LIB3051) cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) endosperm tissue. Seeds are germinated on paper towels under laboratory ambient light conditions. At 8, 10 and 14 hours after imbibition, the seed coats are harvested. The endosperm consists of a very thin layer of tissue affixed to the inside of the seed coat. The seed coat and endosperm are frozen immediately after harvest in liquid nitrogen. The harvested tissue is stored at -80°C until RNA preparation.

The Soy60 (LIB3072) cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) drought stressed seed plus pod subtracted from control seed plus pod tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 26°C and the nighttime temperature 21°C and 70% relative humidity. Soil is checked and watered daily to maintain even moisture conditions. At the R3 stage of the plant drought is induced by withholding water. After 3 and 6 days seeds and pods from both drought stressed and

control (watered regularly) plants are collected from the fifth and sixth node and frozen in dry-ice. The harvested tissue is stored at  $-80^{\circ}\text{C}$  until RNA preparation. For subtraction, target cDNA is made from the drought stressed tissue total RNA using the SMART cDNA synthesis system from Clontech (Clontech Laboratories, Palo Alto, California U.S.A.). Driver first strand cDNA is covalently linked to Dynabeads following a protocol similar to that described in the Dynal literature (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The target cDNA is then heat denatured and the second strand trapped using Dynabeads oligo-dT. The target second strand cDNA is then hybridized to the driver cDNA in 400  $\mu$ l 2X SSPE for two rounds of hybridization at  $65^{\circ}\text{C}$  and 20 hours. After each hybridization, the hybridization solution is removed from the system and the hybridized target cDNA removed from the driver by heat denaturation in water. After hybridization, the remaining cDNA is trapped with Dynabeads oligo-dT. The trapped cDNA is then amplified as in previous PCR based libraries and the resulting cDNA ligated into the pSPORT vector (Invitrogen, Carlsbad California U.S.A.).

The Soy61 (LIB3073) cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) jasmonic acid treated seedling subtracted from control tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in a greenhouse. The daytime temperature is approximately  $29.4^{\circ}\text{C}$  and the nighttime temperature  $20^{\circ}\text{C}$ . Soil is checked and watered daily to maintain even moisture conditions. At 9 days post planting, the plantlets are sprayed with either control buffer of 0.1% Tween-20 or jasmonic acid (Sigma J-2500, Sigma, St. Louis, Missouri U.S.A.) at 1 mg/ml in 0.1% Tween-20. Plants are sprayed until runoff and the soil and the stem is soaked with the spraying solution. At 18 hours post application of jasmonic acid, the soybean plantlets appear growth retarded. After 18hours, 24hours and 48 hours post

669670  
2312260

treatment, the cotyledons are removed and the remaining leaf and stem tissue above the soil is harvested and frozen in liquid nitrogen. The harvested tissue is stored at -80°C until RNA preparation. To make RNA, the three sample timepoints were combined and ground. For subtraction, target cDNA is made from the jasmonic acid treated tissue total RNA using the SMART cDNA synthesis system from Clontech (Clontech Laboratories, Palo Alto, California U.S.A.). Driver first strand cDNA is covalently linked to Dynabeads following a protocol similar to that described in the Dynal literature (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The target cDNA is then heat denatured and the second strand trapped using Dynabeads oligo-dT. The target second strand cDNA is then hybridized to the driver cDNA in 400 12X SSPE for two rounds of hybridization at 65°C and 20 hours. After each hybridization, the hybridization solution is removed from the system and the hybridized target cDNA removed from the driver by heat denaturation in water. After hybridization, the remaining cDNA is trapped with Dynabeads oligo-dT. The trapped cDNA is then amplified as in previous PCR based libraries and the resulting cDNA ligated into the pSPORT vector (Invitrogen, Carlsbad California U.S.A.). For this library's construction, the eighth fraction of the cDNA size fractionation step was used for ligation.

The Soy62 (LIB3074) cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) jasmonic acid treated seedling subtracted from control tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in a greenhouse. The daytime temperature is approximately 29.4°C and the nighttime temperature 20°C. Soil is checked and watered daily to maintain even moisture conditions. At 9 days post planting, the plantlets are sprayed with either control buffer of 0.1% Tween-20 or jasmonic acid (Sigma J-2500, Sigma, St.

Loius, Missouri U.S.A.) at 1 mg/ml in 0.1% Tween-20. Plants are sprayed until runoff and the soil and the stem is soaked with the spraying solution. At 18 hours post application of jasmonic acid, the soybean plantlets appear growth retarded. After 18 hours, 24 hours and 48 hours post treatment, the cotyledons are removed and the remaining leaf and stem tissue above the soil is harvested and frozen in liquid nitrogen. The harvested tissue is stored at -80°C until RNA preparation. To make RNA, the three sample timepoints were combined and ground. For subtraction, target cDNA is made from the jasmonic acid treated tissue total RNA using the SMART cDNA synthesis system from Clontech (Clontech Laboratories, Palo Alto, California U.S.A.). Driver first strand cDNA is covalently linked to Dynabeads following a protocol similar to that described in the Dynal literature (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The target cDNA is then heat denatured and the second strand trapped using Dynabeads oligo-dT. The target second strand cDNA is then hybridized to the driver cDNA in 400 12X SSPE for two rounds of hybridization at 65°C and 20 hours. After each hybridization, the hybridization solution is removed from the system and the hybridized target cDNA removed from the driver by heat denaturation in water. After hybridization, the remaining cDNA is trapped with Dynabeads oligo-dT. The trapped cDNA is then amplified as in previous PCR based libraries and the resulting cDNA ligated into the pSPORT vector (Invitrogen, Carlsbad California U.S.A.). For this library's construction, the ninth fraction of the cDNA size fractionation step was used for ligation.

The Soy65 (LIB3107) 07cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) drought-stressed abscission zone tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in an environmental chamber under 12hr daytime/12hr

nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature 24°C. Soil is checked and watered daily to maintain even moisture conditions. Plants are irrigated with 15-16-17 Peter's Mix. At the R3 stage of development, drought is imposed by withholding water. At 3, 4, 5 and 6 days, tissue is harvested and wilting is not obvious until the fourth day. Abscission layers from reproductive organs are harvested by cutting less than one millimeter proximal and distal to the layer and immediately frozen in liquid nitrogen. The harvested tissue is stored at -80°C until RNA preparation.

The Soy66 (LIB3109) cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) non-drought stressed abscission zone tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature approximately 24°C. Soil is checked and watered daily to maintain even moisture conditions. Plants are irrigated with 15-16-17 Peter's Mix. At 3, 4, 5 and 6 days, control abscission layer tissue is harvested. Abscission layers from reproductive organs are harvested by cutting less than one millimeter proximal and distal to the layer and immediately frozen in liquid nitrogen. The harvested tissue is stored at -80°C until RNA preparation.

Soy67 (LIB3065) cDNA library is prepared from equal amounts tissue harvested from SOYMON007, SOYMON015 and SOYMON020 prepared tissue. Single stranded and double stranded DNA representing approximately  $1 \times 10^6$  colony forming units are isolated using standard protocols. RNA, complementary to the single stranded DNA, is synthesized using the double stranded DNA as a template. Biotinylated dATP is incorporated into the RNA during the synthesis reaction. The single stranded DNA is mixed with the biotinylated RNA in a 1:10 molar

ratio) and allowed to hybridize. DNA-RNA hybrids are captured on Dynabeads M280 streptavidin (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The dynabeads with captured hybrids are collected with a magnet. Captured hybrids are eluted with water.

Soy68 (LIB3052) cDNA library is prepared from equal amounts tissue harvested from SOYMON007, SOYMON015 and SOYMON020 prepared tissue. Single stranded and double stranded DNA representing approximately  $1 \times 10^6$  colony forming units are isolated using standard protocols. RNA, complementary to the single stranded DNA, is synthesized using the double stranded DNA as a template. Biotinylated dATP is incorporated into the RNA during the synthesis reaction. The single stranded DNA is mixed with the biotinylated RNA in a 1:10 molar ratio) and allowed to hybridize. DNA-RNA hybrids are captured on Dynabeads M280 streptavidin (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The dynabeads with captured hybrids are collected with a magnet. Captured hybrids are eluted with water.

Soy69 (LIB3053) cDNA library is generated from soybean cultivars Cristalina (USDA Soybean Germplasm Collection, Urbana, Illinois U.S.A.) and FT108 (Monsoy, Brazil) (tropical germ plasma) normalized leaf tissue. Leaves are harvested from plants grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature approximately 24°C. Soil is checked and watered daily to maintain even moisture conditions. Approximately 30g of leaves are harvested from the 4<sup>th</sup> node of each of the Cristalina and FT108 cultivars and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation. Single stranded and double stranded DNA representing approximately  $1 \times 10^6$  colony forming units are isolated using standard protocols. RNA, complementary to the single stranded DNA, is synthesized using the double stranded DNA as a template. Biotinylated dATP is incorporated into the RNA during the



synthesis reaction. The single stranded DNA is mixed with the biotinylated RNA in a 1:10 molar ratio) and allowed to hybridize. DNA-RNA hybrids are captured on Dynabeads M280 streptavidin (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The dynabeads with captured hybrids are collected with a magnet. The non-hybridized single stranded molecules remaining after hybrid capture are converted to double stranded form and represent the primary normalized library.

Soy70 (LIB3055) cDNA library is generated from soybean cultivars Cristalina (USDA Soybean Germplasm Collection, Urbana, Illinois U.S.A.) and FT108 (Monsoy, Brazil) (tropical germ plasma) leaf tissue. Leaves are harvested from plants grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature approximately 24°C. Soil is checked and watered daily to maintain even moisture conditions. Approximately 30g of leaves are harvested from the 4<sup>th</sup> node of each of the Cristalina and FT108 cultivars and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

Soy71 (LIB3056) cDNA library is generated from soybean cultivars Cristalina and FT108 (tropical germ plasma) root tissue. Roots are harvested from plants grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 29°C and the nighttime temperature approximately 24°C. Soil is checked and watered daily to maintain even moisture conditions. Approximately 50g and 56g of roots are harvested from each of the Cristalina and FT108 cultivars and immediately frozen in dry ice. The harvested tissue is then stored at -80°C until RNA preparation.

Soy72 (LIB3093) cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) drought stressed leaf control tissue. Seeds

are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in an environmental chamber under 12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 26°C and the nighttime temperature 21°C and 70% relative humidity. Soil is checked and watered daily to maintain even moisture conditions. At the R3 stage of the plant drought is induced by withholding water. After 3 and 6 days seeds and pods from both drought stressed and control (watered regularly) plants are collected from the fifth and sixth node and frozen in dry-ice. The harvested tissue is stored at -80°C until RNA preparation. For subtraction, target cDNA is made from the drought stressed tissue total RNA using the SMART cDNA synthesis system from Clontech (Clontech Laboratories, Palo Alto, California U.S.A.). Driver first strand cDNA is covalently linked to Dynabeads following a protocol similar to that described in the Dynal literature (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The target cDNA is then heat denatured and the second strand trapped using Dynabeads oligo-dT. The target second strand cDNA is then hybridized to the driver cDNA in 400 12X SSPE for two rounds of hybridization at 65°C and 20 hours. After each hybridization, the hybridization solution is removed from the system and the hybridized target cDNA removed from the driver by heat denaturation in water. After hybridization, the remaining cDNA is trapped with Dynabeads oligo-dT. The trapped cDNA is then amplified as in previous PCR based libraries and the resulting cDNA ligated into the pSPORT vector (Invitrogen, Carlsbad California U.S.A.).

Soy73 (LIB3093) cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) drought stressed leaf subtracted from control tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in an environmental chamber under

12hr daytime/12hr nighttime cycles. The daytime temperature is approximately 26°C and the nighttime temperature 21°C and 70% relative humidity. Soil is checked and watered daily to maintain even moisture conditions. At the R3 stage of the plant drought is induced by withholding water. After 3 and 6 days seeds and pods from both drought stressed and control (watered regularly) plants are collected from the fifth and sixth node and frozen in dry-ice. The harvested tissue is stored at -80°C until RNA preparation. For subtraction, target cDNA is made from the drought stressed tissue total RNA using the SMART cDNA synthesis system from Clontech (Clontech Laboratories, Palo Alto, California U.S.A.). Driver first strand cDNA is covalently linked to Dynabeads following a protocol similar to that described in the Dynal literature (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The target cDNA is then heat denatured and the second strand trapped using Dynabeads oligo-dT. The target second strand cDNA is then hybridized to the driver cDNA in 400 1 2X SSPE for two rounds of hybridization at 65°C and 20 hours. After each hybridization, the hybridization solution is removed from the system and the hybridized target cDNA removed from the driver by heat denaturation in water. After hybridization, the remaining cDNA is trapped with Dynabeads oligo-dT. The trapped cDNA is then amplified as in previous PCR based libraries and the resulting cDNA ligated into the pSPORT vector (Invitrogen, Carlsbad California U.S.A.).

The Soy76 (Lib3106) cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) jasmonic acid and arachidonic treated seedling subtracted from control tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in a greenhouse. The daytime temperature is approximately 29.4°C and the nighttime temperature 20°C. Soil is checked and watered daily to maintain even moisture conditions. At 9 days post planting, the

plantlets are sprayed with either control buffer of 0.1% Tween-20 or jasmonic acid (Sigma J-2500, Sigma, St. Louis, Missouri U.S.A.) at 1 mg/ml in 0.1% Tween-20. Plants are sprayed until runoff and the soil and the stem is soaked with the spraying solution. At 18 hours post application of jasmonic acid, the soybean plantlets appear growth retarded. Arachidonic treated seedlings are sprayed with 1m/ml arachidonic acid in 0.1% Tween-20. After 18hours, 24hours and 48 hours post treatment, the cotyledons are removed and the remaining leaf and stem tissue above the soil is harvested and frozen in liquid nitrogen. The harvested tissue is stored at -80°C until RNA preparation. To make RNA, the three sample timepoints were combined and ground. The RNA from the arachidonic treated seedlings is isolated separately. For subtraction, target cDNA is made from the jasmonic acid treated tissue total RNA using the SMART cDNA synthesis system from Clontech (Clontech Laboratories, Palo Alto, California U.S.A.). Driver first strand cDNA is covalently linked to Dynabeads following a protocol similar to that described in the Dynal literature (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The target cDNA is then heat denatured and the second strand trapped using Dynabeads oligo-dT. The target second strand cDNA is then hybridized to the driver cDNA in 400 1 2X SSPE for two rounds of hybridization at 65°C and 20 hours. After each hybridization, the hybridization solution is removed from the system and the hybridized target cDNA removed from the driver by heat denaturation in water. After hybridization, the remaining cDNA is trapped with Dynabeads oligo-dT. The trapped cDNA is then amplified as in previous PCR based libraries and the resulting cDNA ligated into the pSPORT vector (Invitrogen, Carlsbad California U.S.A.). Fraction 10 of the size fractionated cDNA is ligated into the pSPORT vector (Invitrogen, Carlsbad California U.S.A.) in order to capture some of the smaller transcripts characteristic of antifungal proteins.

Soy77 (LIB3108) cDNA library is generated from soybean cultivar Asgrow 3244 (Asgrow Seed Company, Des Moines, Iowa U.S.A.) jasmonic acid control tissue. Seeds are planted at a depth of approximately 2cm into 2-3 inch peat pots containing Metromix 350 medium and the plants are grown in a greenhouse. The daytime temperature is approximately 29.4°C and the nighttime temperature 20°C. Soil is checked and watered daily to maintain even moisture conditions. At 9 days post planting, the plantlets are sprayed with either control buffer of 0.1% Tween-20 or jasmonic acid (Sigma J-2500, Sigma, St. Louis, Missouri U.S.A.) at 1 mg/ml in 0.1% Tween-20. Plants are sprayed until runoff and the soil and the stem is soaked with the spraying solution. At 18 hours post application of jasmonic acid, the soybean plantlets appear growth retarded. Arachidonic treated seedlings are sprayed with 1m/ml arachidonic acid in 0.1% Tween-20. After 18 hours, 24 hours and 48 hours post treatment, the cotyledons are removed and the remaining leaf and stem tissue above the soil is harvested and frozen in liquid nitrogen. The harvested tissue is stored at -80°C until RNA preparation. To make RNA, the three sample timepoints were combined and ground. The RNA from the arachidonic treated seedlings is isolated separately. For subtraction, target cDNA is made from the jasmonic acid treated tissue total RNA using the SMART cDNA synthesis system from Clontech (Clontech Laboratories, Palo Alto, California U.S.A.). Driver first strand cDNA is covalently linked to Dynabeads following a protocol similar to that described in the Dynal literature (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.). The target cDNA is then heat denatured and the second strand trapped using Dynabeads oligo-dT. The target second strand cDNA is then hybridized to the driver cDNA in 400 l 2X SSPE for two rounds of hybridization at 65°C and 20 hours. After each hybridization, the hybridization solution is removed from the system and the hybridized target cDNA removed from the driver by heat denaturation in water. After

hybridization, the remaining cDNA is trapped with Dynabeads oligo-dT. The trapped cDNA is then amplified as in previous PCR based libraries and the resulting cDNA ligated into the pSPORT vector (Invitrogen, Carlsbad California U.S.A.). Fraction 10 of the size fractionated cDNA is ligated into the pSPORT vector in order to capture some of the smaller transcripts characteristic of antifungal proteins.

The Lib9 cDNA library is prepared from *Arabidopsis thaliana*, Columbia ecotype, leaf tissue. Wild type *Arabidopsis thaliana* seeds are planted in commonly used planting pots and grown in an environmental chamber. Leaf blades were cut with sharp scissors at seven weeks after planting. The tissue was immediately frozen in liquid nitrogen. The harvested tissue is stored at  $-80^{\circ}\text{C}$  until RNA extraction. PolyA mRNA is purified from the total RNA preparation using Dynabeads® Oligo(dT)<sub>25</sub> (DynaL Inc., Lake Success, N.Y.), or equivalent methods. This library was normalized using a PCR-based protocol.

The Lib22 cDNA library is prepared from *Arabidopsis thaliana* Columbia ecotype, root tissue. Wild type *Arabidopsis thaliana* seeds are planted in commonly used planting pots and grown in an environmental chamber. After 5-6 weeks the plants are in the reproductive growth phase. Stems are bolting from the base of the plants. After 7 weeks, more stems, floral buds appear, and a few flowers are starting to open. The 7-week old plants are rinsed intensively by tope water remove dirt from the roots, and blotted by paper towel. The tissues are immediately frozen in liquid nitrogen. The harvested tissue is stored at  $-80^{\circ}\text{C}$  until RNA preparation.

The Lib23 cDNA library is prepared from *Arabidopsis thaliana*, Columbia ecotype, stem tissue. Wild type *Arabidopsis thaliana* seeds are planted in commonly used planting pots and grown in an environmental chamber. Stems were collected seven to eight weeks after planting

by cutting the stems from the base and cutting the top of the plant to remove the floral tissue. The tissue was immediately frozen in liquid nitrogen and stored at -80°C until total RNA extraction. PolyA mRNA is purified from the total RNA preparation using Dynabeads® Oligo(dT)<sub>25</sub> (DynaL Inc., Lake Success, N.Y.), or equivalent methods. This library was normalized using a PCR-based protocol.

The Lib24 cDNA library is prepared from *Arabidopsis thaliana*, Columbia ecotype, flower bud tissue. Wild type *Arabidopsis thaliana* seeds are planted in commonly used planting pots and grown in an environmental chamber. Flower buds are green and unopened and harvested about seven weeks after planting. The tissue is immediately frozen in liquid nitrogen. The harvested tissue is stored at -80°C until total RNA extraction. PolyA mRNA is purified from the total RNA preparation using Dynabeads® Oligo(dT)<sub>25</sub> (DynaL Inc., Lake Success, N.Y.), or equivalent methods. This library was normalized using a PCR-based protocol.

The Lib25 cDNA library is prepared from *Arabidopsis thaliana*, Columbia ecotype, open flower tissue. Wild type *Arabidopsis thaliana* seeds are planted in commonly used planting pots and grown in an environmental chamber. Flowers are completely opened with all parts of floral structure observable, but no siliques are appearing. The tissue was immediately frozen in liquid nitrogen and stored at -80°C until total RNA extraction. PolyA mRNA is purified from the total RNA preparation using Dynabeads® Oligo(dT)<sub>25</sub> (DynaL Inc., Lake Success, N.Y.), or equivalent methods. This library was normalized using a PCR-based protocol.

The Lib35 cDNA library of the present invention, was prepared from *Arabidopsis thaliana* Columbia ecotype leaf tissue. Wild type *Arabidopsis thaliana* seeds are planted in commonly used planting pots and grown in an environmental chamber. After 5-6 weeks the

plants are in the reproductive growth phase. Stems are bolting from the base of the plants. After 7 weeks, more stems and floral buds appeared and a few flowers were starting to open. Leaf blades were collected by cutting with sharp scissors. The tissues were immediately frozen in liquid nitrogen and stored at -80°C until use. PolyA mRNA is purified from the total RNA preparation using Dynabeads® Oligo(dT)<sub>25</sub> (DynaL Inc., Lake Success, N.Y.), or equivalent methods. This library was normalized using a PCR-based protocol.

The Lib146 cDNA library is prepared from *Arabidopsis thaliana*, Columbia ecotype, immature seed tissue. Wild type *Arabidopsis thaliana* seeds are planted in commonly used planting pots and grown in an environmental chamber. At approximately 7-8 weeks of age, the seeds are harvested. The seeds ranged in maturity from the smallest seeds that could be dissected from silques to just before starting to turn yellow in color. The tissue is immediately frozen in liquid nitrogen. The harvested tissue is stored at -80°C until RNA extraction. PolyA mRNA is purified from the total RNA preparation using Dynabeads® Oligo(dT)<sub>25</sub> (DynaL Inc., Lake Success, N.Y.), or equivalent methods. This library is normalized using a PCR-based protocol.

The Lib3032 (Lib80) cDNA libraries are generated from *Brassica napus* seeds harvested 30 days after pollination. The cDNA libraries are constructed using the SuperScript Plasmid system for cDNA synthesis and plasmid cloning (Life Technologies, Gaithersburg, Maryland U.S.A.) according to the manufacturers protocol with the following modification: 40 micrograms of total RNA is used as the starting material for cDNA synthesis, and first strand cDNA synthesis is carried out at 45°C.

The Lib3034 (Lib82) cDNA libraries are generated from *Brassica napus* seeds harvested 15 and 18 days after pollination. The cDNA libraries are constructed using the SuperScript



Plasmid system for cDNA synthesis and plasmid cloning (Life Technologies, Gaithersburg, Maryland U.S.A.) according to the manufacturers protocol with the following modification: 40 micrograms of total RNA is used as the starting material for cDNA synthesis, and first strand cDNA synthesis was carried out at 45°C.

The Lib3099 cDNA library is generated by a subtraction procedure. The library contains cDNAs whose abundance is enriched in the *Brassica napus* 15 and 18 day after pollination seed tissues when compared to *Brassica* leaf tissues. The cDNA synthesis is performed on *Brassica* leaf RNA and *Brassica* RNA isolated from seeds harvested 15 and 18 days after pollination using a Smart PCR cDNA synthesis kit according to the manufacturers protocol (Clontech, Palo Alto, California U.S.A.). The subtracted cDNA is generated using the Clontech PCR-Select subtraction kit according to the manufacturers protocol (Clontech, Palo Alto, California U.S.A.). The subtracted cDNA was cloned into plasmid vector pCR2.1 according to the manufacturers protocol (Invitrogen, Carlsbad, California U.S.A.).

The Lib3033 (Lib81) cDNA libraries are generated from from the *Schizochytrium* species cells. The *Schizochytrium* species cells are grown in liquid media until saturation. The culture is centrifuged to pellet the cells, the medium is decanted off, and pellet immediately frozen in liquid nitrogen. Wax esters are produced under such dark, anaerobic, rich-medium conditions. High wax production by the cultures is verified by microscopy (fluorescein staining of wax bodies) and by lipid extraction/TLC/GC. The harvested cells are stored at -80°C until RNA preparation. RNA is prepared from the frozen *Euglena* cell pellet as follows. The pellet is pulverized to a powder in liquid nitrogen using a mortar and pestle. The powder is transferred to tubes containing 6 ml each of lysis buffer (100 mM Tris, pH 8, 0.6 M NaCl, 10 mM EDTA, and 4% (w/v) SDS) and buffered phenol, vortexed, and disrupted with a Polytron. The mixture is

centrifuged 20 min at 10,000xg in Corex glass tubes to separate the phases. 5 ml of the upper phase is removed, vortexed with 5 ml fresh phenol, and centrifuged. The upper phase is removed and the RNA is precipitated overnight at 4°C by adding 1.5 volumes of 4 M LiCl. The RNA is further purified on Rneasy columns according to the manufacturers protocol (Qiagen, Valencia, California U.S.A.). The cDNA library is constructed using the SuperScript Plasmid system for cDNA synthesis and plasmid cloning (Life Technologies, Gaithersburg, Maryland U.S.A.) according to the manufacturers protocol with the following modification: 40 micrograms of total RNA was used as the starting material for cDNA synthesis, and first strand cDNA synthesis was carried out at 45°C.

The Lib47 cDNA library is generated from *Euglena gracilis* strain 753 (ATTC No. 30285, ATCC Manasas, Virginia U.S.A.) grown in liquid culture. A liquid culture is inoculated with 1/10 volume of a previously-grown saturated culture, and the new culture for 4 days under near-anaerobic conditions (near-anaerobic cultures are not agitated, just gently swirled once a day) in the dark in 2X Beef (10 g/l bacto peptone, 4 g/l yeast extract, 2 g/l beef extract, 6 g/l glucose). The culture is then centrifuged to pellet the cells, the medium is decanted off, and pellet immediately frozen in liquid nitrogen. Wax esters are produced under such dark, anaerobic, rich-medium conditions. High wax production by the cultures is verified by microscopy (fluorescein staining of wax bodies) and by lipid extraction/TLC/GC. The harvested cells are stored at -80°C until RNA preparation. RNA is prepared from the frozen *Euglena* cell pellet as follows. The pellet is pulverized to a powder in liquid nitrogen using a mortar and pestle. The powder is transferred to tubes containing 6 ml each of lysis buffer (100 mM Tris, pH 8, 0.6 M NaCl, 10 mM EDTA, and 4% (w/v) SDS) and buffered phenol, vortexed, and disrupted with a Polytron. The mixture is centrifuged 20 min at 10,000xg in Corex glass tubes to separate

the phases. 5 ml of the upper phase is removed, vortexed with 5 ml fresh phenol, and centrifuged. The upper phase is removed and the RNA is precipitated overnight at 4°C by adding 1.5 volumes of 4 M LiCl. The RNA is further purified on Rneasy columns according to the manufacturers protocol (Qiagen, Valencia, California U.S.A.). The cDNA library is constructed using the SuperScript Plasmid system for cDNA synthesis and plasmid cloning (Life Technologies, Gaithersburg, Maryland U.S.A.) according to the manufacturers protocol with the following modification: 40 micrograms of total RNA was used as the starting material for cDNA synthesis, and first strand cDNA synthesis was carried out at 45°C.

The Lib44 cDNA library is generated from *Phaeodactylum tricornatum* grown in modified Jones medium for 3 days. The cells were harvested by centrifugation and the resulting pellet frozen immediately in liquid nitrogen. The harvested cells are stored at -80°C until RNA preparation. RNA is prepared from the frozen *Phaeodactylum* cell pellet as follows. The pellet is pulverized to a powder in liquid nitrogen using a mortar and pestle. The powder is transferred to tubes containing 6 ml each of lysis buffer (100 mM Tris, pH 8, 0.6 M NaCl, 10 mM EDTA, and 4% (w/v) SDS) and buffered phenol, vortexed, and disrupted with a Polytron. The mixture is centrifuged 20 min at 10,000xg in Corex glass tubes to separate the phases. 5 ml of the upper phase is removed, vortexed with 5 ml fresh phenol, and centrifuged. The upper phase is removed and the RNA is precipitated overnight at 4°C by adding 1.5 volumes of 4 M LiCl. The RNA is further purified on Rneasy columns according to the manufacturers protocol (Qiagen, Valencia, California U.S.A.). The cDNA library is constructed using the SuperScript Plasmid system for cDNA synthesis and plasmid cloning (Life Technologies, Gaithersburg, Maryland U.S.A.) according to the manufacturers protocol with the following modification: 40 micrograms of total

RNA was used as the starting material for cDNA synthesis, and first strand cDNA synthesis was carried out at 45 degrees centigrade.

The LIB3036 genomic library is generated from *Mycobacterium neoaurum* US52 (ATCC No. 23072, ATCC, Manassas, Virginia U.S.A.) cells. *Mycobacterium neoaurum* US52 is a gram-positive Actinomycete bacterium. *Mycobacterium neoaurum* US52 is genetically related to *Mycobacterium tuberculosis*, but there is no reason to believe that it is a primary pathogen. It normally is saprophytic, i.e. it lives in soil and gets nutrients from decaying matter. Genomic DNA obtained from *Mycobacterium neoaurum* US52 is digested for various times with the restriction enzyme Sau3A. The DNA fractions are size-separated on an agarose gel, and the first fraction wherein most of the partially-digested fragments are about 10 kB is used to isolated fragments in the range of 2-3 kB. For LIB3036, the 2-3 kB DNA is cloned into vector pRY401 (Invitrogen, Carlsbad, California U.S.A.). The vector pZERO-2 (Invitrogen, Carlsbad, California U.S.A.). is used for the construction of LIB3104.

The stored RNA is purified using Trizol reagent from Life Technologies (Gibco BRL, Life Technologies, Gaithersburg, Maryland U.S.A.), essentially as recommended by the manufacturer. Poly A+ RNA (mRNA) is purified using magnetic oligo dT beads essentially as recommended by the manufacturer (Dynabeads, Dynal Corporation, Lake Success, New York U.S.A.).

Construction of plant cDNA libraries is well-known in the art and a number of cloning strategies exist. A number of cDNA library construction kits are commercially available. The Superscript™ Plasmid System for cDNA synthesis and Plasmid Cloning (Gibco BRL, Life

Technologies, Gaithersburg, Maryland U.S.A.) is used, following the conditions suggested by the manufacturer.

Normalized libraries are made using essentially the Soares procedure (Soares *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 91:9228-9232 (1994), the entirety of which is herein incorporated by reference). This approach is designed to reduce the initial 10,000-fold variation in individual cDNA frequencies to achieve abundances within one order of magnitude while maintaining the overall sequence complexity of the library. In the normalization process, the prevalence of high-abundance cDNA clones decreases dramatically, clones with mid-level abundance are relatively unaffected and clones for rare transcripts are effectively increased in abundance.

### **Example 2**

The cDNA libraries are plated on LB agar containing the appropriate antibiotics for selection and incubated at 37° for a sufficient time to allow the growth of individual colonies. Single colonies are individually placed in each well of a 96-well microtiter plates containing LB liquid including the selective antibiotics. The plates are incubated overnight at approximately 37°C with gentle shaking to promote growth of the cultures. The plasmid DNA is isolated from each clone using Qiaprep plasmid isolation kits, using the conditions recommended by the manufacturer (Qiagen Inc., Santa Clara, California U.S.A.).

Template plasmid DNA clones are used for subsequent sequencing. For sequencing, the ABI PRISM dRhodamine Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq® DNA Polymerase, FS, is used (PE Applied Biosystems, Foster City, California U.S.A.).

### **Example 3**

Nucleic acid sequences that encode for the following proteins: triose phosphate isomerase, fructose 1,6-bisphosphate aldolase, fructose 1,6-bisphosphate, fructose 6-phosphate 2-kinase, phosphoglucosomerase, vacuolar H<sup>+</sup> translocating-pyrophosphatase, pyrophosphate-dependent fructose-6-phosphate phosphotransferase, invertase, sucrose synthase, hexokinase, fructokinase, NDP-kinase, glucose-6-phosphate 1-dehydrogenase, phosphoglucomutase and UDP-glucose pyrophosphorylase are identified from the Monsanto EST PhytoSeq database using TBLASTN (default values)(TBLASTN compares a protein query against the six reading frames of a nucleic acid sequence). Matches found with BLAST P values equal or less than 0.001 (probability) or BLAST Score of equal or greater than 90 are classified as hits. If the program used to determine the hit is HMMSW then the score refers to HMMSW score.

In addition, the GenBank database is searched with BLASTN and BLASTX (default values) using ESTs as queries. EST that pass the hit probability threshold of  $10e^{-8}$  for the following enzymes are combined with the hits generated by using TBLASTN (described above) and classified by enzyme (see Table A below).

A cluster refers to a set of overlapping clones in the PhytoSeq database. Such an overlapping relationship among clones is designated as a “cluster” when BLAST scores from pairwise sequence comparisons of the member clones meets a predetermined minimum value or product score of 50 or more (Product Score = (BLAST SCORE x Percentage Identity)/(5 x minimum [length (Seq1), length (Seq2)]))

Since clusters are formed on the basis of single-linkage relationships, it is possible for two non-overlapping clones to be members of the same cluster if, for instance, they both overlap a third clone with at least the predetermined minimum BLAST score (stringency). A cluster ID

is arbitrarily assigned to all of those clones which belong to the same cluster at a given stringency and a particular clone will belong to only one cluster at a given stringency. If a cluster contains only a single clone (a “singleton”), then the cluster ID number will be negative, with an absolute value equal to the clone ID number of its single member. Clones grouped in a cluster in most cases represent a contiguous sequence.

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.26.000000>; this version posted March 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TABLE A\*

MAIZE TRIOSE PHOSPHATE ISOMERASE								
Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
1	-700019675	700019675H1	SATMON001	g546735	BLASTX	134	1e-11	78
2	-700073894	700073894H1	SATMON007	g609261	BLASTN	257	1e-10	84
3	-700167260	700167260H1	SATMON013	g609261	BLASTN	644	1e-44	79
4	-700380595	700380595H1	SATMON021	g609261	BLASTN	1121	1e-84	87
5	-700449667	700449667H1	SATMON028	g217973	BLASTN	204	1e-18	93
6	-700449720	700449720H2	SATMON028	g217973	BLASTN	216	1e-18	88
7	-700570661	700570661H1	SATMON030	g168647	BLASTX	131	1e-11	88
8	-700616770	700616770H1	SATMON033	g407525	BLASTX	149	1e-13	83
9	-701170944	701170944H1	SATMONN05	g217921	BLASTX	188	1e-20	53
10	11337	700337974H1	SATMON020	g256119	BLASTN	535	1e-61	78
11	11337	700027829H1	SATMON003	g256119	BLASTN	726	1e-51	80
12	126	700050046H1	SATMON003	g1785947	BLASTN	440	1e-26	92
13	282	700077320H1	SATMON007	g217973	BLASTN	666	1e-108	97
14	282	700104541H1	SATMON010	g217973	BLASTN	631	1e-106	97
15	282	700047476H1	SATMON003	g217973	BLASTN	648	1e-105	97
16	282	700211559H1	SATMON016	g217973	BLASTN	525	1e-104	97
17	282	700073553H1	SATMON007	g217973	BLASTN	981	1e-103	98
18	282	700613011H1	SATMON033	g217973	BLASTN	552	1e-102	98
19	282	700352119H1	SATMON023	g217973	BLASTN	666	1e-101	97
20	282	700088148H1	SATMON011	g217973	BLASTN	666	1e-100	98
21	282	700351626H1	SATMON023	g217973	BLASTN	401	1e-99	98
22	282	700240096H1	SATMON010	g217973	BLASTN	666	1e-98	97
23	282	700083660H1	SATMON011	g217973	BLASTN	666	1e-97	99
24	282	700208721H1	SATMON016	g217973	BLASTN	497	1e-96	98
25	282	700203144H1	SATMON003	g217973	BLASTN	511	1e-96	96
26	282	700430425H1	SATMONN01	g217973	BLASTN	666	1e-96	98
27	282	700206091H1	SATMON003	g217973	BLASTN	497	1e-94	97
28	282	700077017H1	SATMON007	g217973	BLASTN	614	1e-93	93
29	282	700618792H1	SATMON034	g217973	BLASTN	546	1e-92	96
30	282	700572532H1	SATMON030	g407524	BLASTN	1212	1e-92	84
31	282	700106512H1	SATMON010	g217973	BLASTN	632	1e-91	97
32	282	700195031H1	SATMON014	g217973	BLASTN	471	1e-90	97
33	282	700168131H1	SATMON013	g217973	BLASTN	497	1e-89	98
34	282	700197039H1	SATMON014	g217973	BLASTN	546	1e-89	98
35	282	700572688H1	SATMON030	g169820	BLASTN	1114	1e-89	85
36	282	700021313H1	SATMON001	g217973	BLASTN	913	1e-87	97
37	282	700452417H1	SATMON028	g217973	BLASTN	425	1e-86	95
38	282	700346119H1	SATMON021	g217973	BLASTN	444	1e-86	96
39	282	700082359H1	SATMON011	g217973	BLASTN	542	1e-86	93
40	282	700240042H1	SATMON010	g217973	BLASTN	596	1e-86	97
41	282	700030064H1	SATMON003	g217973	BLASTN	587	1e-85	94
42	282	700615185H1	SATMON033	g217973	BLASTN	430	1e-84	98
43	282	700196125H1	SATMON014	g217973	BLASTN	581	1e-84	100
44	282	700243429H1	SATMON010	g217973	BLASTN	632	1e-84	97
45	282	700474112H1	SATMON025	g217973	BLASTN	570	1e-83	98
46	282	700572282H1	SATMON030	g407524	BLASTN	838	1e-83	82
47	282	700622238H1	SATMON034	g169820	BLASTN	917	1e-80	86
48	282	700095609H1	SATMON008	g169820	BLASTN	1067	1e-80	82



49	282	700218886H1	SATMON011	g217973	BLASTN	551	1e-79	93
50	282	700018688H1	SATMON001	g217973	BLASTN	1066	1e-79	99
51	282	700049775H1	SATMON003	g217973	BLASTN	362	1e-78	91
52	282	700575972H1	SATMON030	g169820	BLASTN	894	1e-78	79
53	282	700215519H1	SATMON016	g217973	BLASTN	497	1e-76	97
54	282	700161120H1	SATMON012	g217973	BLASTN	622	1e-76	98
55	282	700581760H1	SATMON031	g217973	BLASTN	533	1e-75	90
56	282	700104672H1	SATMON010	g169820	BLASTN	1012	1e-75	83
57	282	700346053H1	SATMON021	g169820	BLASTN	1012	1e-75	83
58	282	701166592H1	SATMONN04	g217973	BLASTN	661	1e-74	95
59	282	700968667H1	SATMONN04	g217973	BLASTN	497	1e-73	92
60	282	700205627H1	SATMON003	g217973	BLASTN	666	1e-73	99
61	282	700029005H1	SATMON003	g169820	BLASTN	979	1e-72	85
62	282	700476479H1	SATMON025	g169820	BLASTN	554	1e-71	84
63	282	700050148H1	SATMON003	g169820	BLASTN	608	1e-70	83
64	282	700259846H1	SATMON017	g217973	BLASTN	283	1e-69	94
65	282	700344093H1	SATMON021	g169820	BLASTN	934	1e-69	83
66	282	700082327H1	SATMON011	g169820	BLASTN	943	1e-69	85
67	282	700020156H1	SATMON001	g217973	BLASTN	420	1e-68	99
68	282	700577714H1	SATMON031	g169820	BLASTN	928	1e-68	85
69	282	700104904H1	SATMON010	g169820	BLASTN	913	1e-67	84
70	282	700104685H1	SATMON010	g169820	BLASTN	897	1e-66	84
71	282	700053463H1	SATMON009	g169820	BLASTN	907	1e-66	85
72	282	700171639H1	SATMON013	g217973	BLASTN	401	1e-65	98
73	282	700574233H1	SATMON030	g169820	BLASTN	651	1e-65	83
74	282	700262653H1	SATMON017	g169820	BLASTN	877	1e-64	84
75	282	700456738H1	SATMON029	g169820	BLASTN	877	1e-64	84
76	282	700611806H1	SATMON022	g169820	BLASTN	877	1e-64	83
77	282	700381177H1	SATMON023	g169820	BLASTN	884	1e-64	84
78	282	700103347H1	SATMON010	g169820	BLASTN	861	1e-63	84
79	282	700103605H1	SATMON010	g169820	BLASTN	868	1e-63	84
80	282	700578536H1	SATMON031	g169820	BLASTN	856	1e-62	84
81	282	700258606H1	SATMON017	g169820	BLASTN	807	1e-61	83
82	282	700335703H1	SATMON019	g217973	BLASTN	376	1e-60	90
83	282	700351044H1	SATMON023	g169820	BLASTN	471	1e-59	83
84	282	700346364H1	SATMON021	g169820	BLASTN	813	1e-59	85
85	282	700619037H1	SATMON034	g169820	BLASTN	814	1e-59	84
86	282	700465160H1	SATMON025	g169820	BLASTN	751	1e-57	84
87	282	700235687H1	SATMON010	g169820	BLASTN	791	1e-57	82
88	282	700105645H1	SATMON010	g169820	BLASTN	793	1e-57	83
89	282	700082237H1	SATMON011	g169820	BLASTN	793	1e-57	84
90	282	700261906H1	SATMON017	g169820	BLASTN	796	1e-57	83
91	282	700456154H1	SATMON029	g169820	BLASTN	799	1e-57	84
92	282	700047696H1	SATMON003	g169820	BLASTN	561	1e-56	83
93	282	700449905H1	SATMON028	g169820	BLASTN	788	1e-56	84
94	282	700336106H1	SATMON019	g217973	BLASTN	325	1e-55	92
95	282	700381867H1	SATMON023	g2529386	BLASTN	422	1e-55	97
96	282	700051335H1	SATMON003	g169820	BLASTN	608	1e-55	83
97	282	700050988H1	SATMON003	g169820	BLASTN	768	1e-55	86
98	282	700029471H1	SATMON003	g169820	BLASTN	772	1e-55	84
99	282	700106806H1	SATMON010	g169820	BLASTN	773	1e-55	84
100	282	700071749H1	SATMON007	g217973	BLASTN	362	1e-54	85
101	282	700207607H1	SATMON016	g217973	BLASTN	362	1e-54	85
102	282	700573465H2	SATMON030	g169820	BLASTN	753	1e-54	86

103	282	700220908H1	SATMON011	g169820	BLASTN	758	1e-54	84
104	282	700467719H1	SATMON025	g169820	BLASTN	761	1e-54	85
105	282	700456018H1	SATMON029	g169820	BLASTN	764	1e-54	81
106	282	700453767H1	SATMON029	g217973	BLASTN	296	1e-52	94
107	282	700026118H1	SATMON003	g217973	BLASTN	341	1e-52	93
108	282	700026760H1	SATMON003	g217973	BLASTN	421	1e-52	99
109	282	700029525H1	SATMON003	g169820	BLASTN	738	1e-52	85
110	282	700457972H1	SATMON029	g169820	BLASTN	723	1e-51	85
111	282	700455866H1	SATMON029	g169820	BLASTN	726	1e-51	84
112	282	700165290H1	SATMON013	g169820	BLASTN	726	1e-51	84
113	282	700351190H1	SATMON023	g169820	BLASTN	672	1e-50	81
114	282	700154095H1	SATMON007	g169820	BLASTN	696	1e-49	84
115	282	700450438H1	SATMON028	g217973	BLASTN	430	1e-48	99
116	282	700044892H1	SATMON004	g169820	BLASTN	683	1e-48	85
117	282	700185095H1	SATMON014	g169820	BLASTN	673	1e-47	84
118	282	700575506H1	SATMON030	g169820	BLASTN	680	1e-47	83
119	282	700161966H1	SATMON012	g217973	BLASTN	335	1e-46	98
120	282	700343401H1	SATMON021	g169820	BLASTN	426	1e-45	77
121	282	700152354H1	SATMON007	g169820	BLASTN	653	1e-45	84
122	282	701164924H1	SATMONN04	g169820	BLASTN	397	1e-44	84
123	282	700346896H1	SATMON021	g169820	BLASTN	496	1e-42	84
124	282	700210157H1	SATMON016	g169820	BLASTN	617	1e-42	84
125	282	700383103H1	SATMON024	g169820	BLASTN	531	1e-41	84
126	282	701158829H1	SATMONN04	g407524	BLASTN	549	1e-40	80
127	282	700619883H1	SATMON034	g217973	BLASTN	325	1e-38	99
128	282	700168219H1	SATMON013	g169820	BLASTN	540	1e-36	83
129	282	700155210H1	SATMON007	g169820	BLASTN	545	1e-36	83
130	282	700334861H1	SATMON019	g169820	BLASTN	484	1e-31	82
131	282	700355663H1	SATMON024	g217973	BLASTN	213	1e-30	88
132	282	700074764H1	SATMON007	g546734	BLASTN	387	1e-27	84
133	282	700621934H1	SATMON034	g217973	BLASTN	430	1e-26	100
134	282	700802084H1	SATMON036	g217973	BLASTN	270	1e-24	98
135	3039	700620444H1	SATMON034	g1785947	BLASTN	473	1e-56	75
136	3039	700356205H1	SATMON024	g1785947	BLASTN	332	1e-32	72
137	3039	700215549H1	SATMON016	g414549	BLASTN	443	1e-26	72
138	3039	700620318H1	SATMON034	g556171	BLASTX	214	1e-25	79
139	3039	700028742H1	SATMON003	g556171	BLASTX	156	1e-20	86
140	3039	700150060H1	SATMON007	g556171	BLASTX	181	1e-17	89
141	3039	700448477H1	SATMON027	g556171	BLASTX	137	1e-12	85
142	3039	700336489H1	SATMON019	g556171	BLASTX	126	1e-10	81
143	3414	700099709H1	SATMON009	g609261	BLASTN	600	1e-49	84
144	3414	700075837H1	SATMON007	g609261	BLASTN	494	1e-41	84
145	3414	700045678H1	SATMON004	g609261	BLASTN	340	1e-29	73
146	3414	700097852H1	SATMON009	g609261	BLASTN	436	1e-27	84
147	3414	700053342H1	SATMON009	g609261	BLASTN	346	1e-25	73
148	3414	700041954H1	SATMON004	g609261	BLASTN	340	1e-24	82
149	3414	700217471H1	SATMON016	g609261	BLASTN	265	1e-21	71
150	3414	700264437H1	SATMON017	g609261	BLASTN	231	1e-17	69
151	3414	700218371H1	SATMON016	g609261	BLASTN	156	1e-10	68
152	5593	700381686H1	SATMON023	g609261	BLASTN	534	1e-44	89
153	5593	700356082H1	SATMON024	g609261	BLASTN	246	1e-24	90
154	5593	700622077H1	SATMON034	g609261	BLASTN	292	1e-20	86
155	5593	700470822H1	SATMON025	g609262	BLASTX	134	1e-11	79
156	6525	700083139H1	SATMON011	g256119	BLASTN	880	1e-64	76

157	6525	700205474H1	SATMON003	g169820	BLASTN	849	1e-62	77
158	6991	700336856H1	SATMON019	g609261	BLASTN	1131	1e-85	85
159	6991	700042717H1	SATMON004	g609261	BLASTN	1028	1e-76	85
160	6991	700379491H1	SATMON020	g609261	BLASTN	995	1e-74	81
161	6991	700156635H1	SATMON012	g609261	BLASTN	877	1e-64	84
162	6991	700046340H1	SATMON004	g609261	BLASTN	852	1e-62	84
163	6991	700081869H1	SATMON011	g609261	BLASTN	266	1e-14	80
164	6991	700426102H1	SATMONN01	g806312	BLASTX	134	1e-13	89
165	7384	700613626H1	SATMON033	g609261	BLASTN	920	1e-87	85
166	7384	700101506H1	SATMON009	g609261	BLASTN	1124	1e-84	85
167	7384	700206445H1	SATMON003	g609261	BLASTN	987	1e-73	79
168	7384	700220160H1	SATMON011	g609261	BLASTN	878	1e-64	85
169	-L1431527	LIB143-004-Q1-E1-C5	LIB143	g217973	BLASTN	290	1e-13	93
170	-L30613868	LIB3061-017-Q1-K1-C9	LIB3061	g217973	BLASTN	182	1e-13	70
171	-L30623620	LIB3062-034-Q1-K1-A8	LIB3062	g609261	BLASTN	599	1e-39	74
172	-L361705	LIB36-021-Q1-E1-E7	LIB36	g609261	BLASTN	266	1e-14	80
173	23992	LIB3062-056-Q1-K1-F9	LIB3062	g1200507	BLASTX	285	1e-64	61
174	282	LIB3067-047-Q1-K1-H2	LIB3067	g217973	BLASTN	1076	1e-164	96
175	282	LIB3067-055-Q1-K1-G8	LIB3067	g217973	BLASTN	1076	1e-133	93
176	282	LIB3067-059-Q1-K1-D10	LIB3067	g169820	BLASTN	1401	1e-115	84
177	282	LIB3067-027-Q1-K1-B10	LIB3067	g407524	BLASTN	995	1e-113	83
178	282	LIB189-032-Q1-E1-H2	LIB189	g217973	BLASTN	629	1e-111	93
179	282	LIB3059-023-Q1-K1-A7	LIB3059	g407524	BLASTN	1436	1e-111	83
180	282	LIB3069-016-Q1-K1-D9	LIB3069	g169820	BLASTN	1301	1e-107	81
181	282	LIB143-006-Q1-E1-A8	LIB143	g169820	BLASTN	1373	1e-105	84
182	282	LIB3068-054-Q1-K1-C11	LIB3068	g169820	BLASTN	1327	1e-102	82
183	282	LIB3067-034-Q1-K1-B7	LIB3067	g407524	BLASTN	1321	1e-101	83
184	282	LIB143-031-Q1-E1-E5	LIB143	g169820	BLASTN	1311	1e-100	84
185	282	LIB3069-055-Q1-K1-H12	LIB3069	g169820	BLASTN	1046	1e-97	75
186	282	LIB3061-027-Q1-K1-A8	LIB3061	g169820	BLASTN	936	1e-96	83
187	282	LIB3078-008-Q1-K1-E5	LIB3078	g169820	BLASTN	1210	1e-92	82
188	282	LIB3066-027-Q1-K1-E1	LIB3066	g407524	BLASTN	1196	1e-91	82
189	282	LIB3067-032-Q1-K1-E5	LIB3067	g169820	BLASTN	1122	1e-84	84

190	282	LIB3078-029-Q1-K1-F7	LIB3078	g169820	BLASTN	827	1e-83	82
191	282	LIB3061-006-Q1-K1-B7	LIB3061	g169820	BLASTN	1091	1e-82	78
192	282	LIB143-048-Q1-E1-F8	LIB143	g169820	BLASTN	644	1e-74	75
193	282	LIB3078-033-Q1-K1-B10	LIB3078	g169820	BLASTN	584	1e-73	79
194	282	LIB3069-046-Q1-K1-C4	LIB3069	g169820	BLASTN	819	1e-59	79
195	282	LIB3061-049-Q1-K1-H2	LIB3061	g169820	BLASTN	587	1e-47	80
196	282	LIB143-029-Q1-E1-G4	LIB143	g169820	BLASTN	679	1e-47	84
197	282	LIB84-027-Q1-E1-E5	LIB84	g169820	BLASTN	613	1e-46	78
198	282	LIB3062-001-Q1-K2-F7	LIB3062	g169820	BLASTN	507	1e-33	80
199	282	LIB3066-014-Q1-K1-H11	LIB3066	g169820	BLASTN	385	1e-25	76
200	29645	LIB3069-014-Q1-K1-C11	LIB3069	g168647	BLASTX	131	1e-27	34
201	29645	LIB3069-013-Q1-K1-C11	LIB3069	g168647	BLASTX	124	1e-24	33
202	3039	LIB3062-045-Q1-K1-F6	LIB3062	g1785947	BLASTN	1119	1e-84	72
203	5593	LIB3067-045-Q1-K1-E5	LIB3067	g609261	BLASTN	702	1e-58	75
204	6991	LIB3059-026-Q1-K1-G9	LIB3059	g609261	BLASTN	1493	1e-115	84
205	6991	LIB3078-049-Q1-K1-E4	LIB3078	g609261	BLASTN	747	1e-55	83
206	7384	LIB3062-034-Q1-K1-A4	LIB3062	g609261	BLASTN	1351	1e-107	85

## MAIZE FRUCTOSE 1,6-BISPHOSPHATE ALDOLASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
207	-700026544	700026544H1	SATMON003	g22144	BLASTN	215	1e-30	88
208	-700073329	700073329H1	SATMON007	g22144	BLASTN	590	1e-89	95
209	-700151987	700151987H1	SATMON007	g22144	BLASTN	212	1e-8	78
210	-700206575	700206575H1	SATMON003	g22144	BLASTN	1009	1e-109	96
211	-700333727	700333727H1	SATMON019	g1217893	BLASTX	154	1e-16	61
212	-700429795	700429795H1	SATMONN01	g1619605	BLASTX	102	1e-16	77
213	-700804137	700804137H1	SATMON036	g22144	BLASTN	742	1e-52	92
214	1182	700449930H1	SATMON028	g22632	BLASTN	856	1e-62	79
215	1182	701185559H1	SATMONN06	g22632	BLASTN	793	1e-57	79
216	1182	700203130H1	SATMON003	g22632	BLASTN	799	1e-57	78
217	1182	700083459H1	SATMON011	g22632	BLASTN	800	1e-57	76
218	1182	700465449H1	SATMON025	g22632	BLASTN	405	1e-50	76
219	1182	701165344H1	SATMONN04	g22632	BLASTN	326	1e-29	78
220	1182	700427538H1	SATMONN01	g438275	BLASTX	96	1e-9	88
221	38	700224356H1	SATMON011	g22144	BLASTN	1290	1e-98	96

222	38	700048169H1	SATMON003	g22144	BLASTN	528	1e-72	98
223	38	700616610H1	SATMON033	g22144	BLASTN	278	1e-31	91
224	38	700355765H1	SATMON024	g20204	BLASTX	141	1e-12	96
225	6547	700194431H1	SATMON014	g2636513	BLASTX	181	1e-17	47
226	6547	700469777H1	SATMON025	g2636513	BLASTX	174	1e-16	48
227	8494	700425929H1	SATMONN01	g927507	BLASTX	67	1e-11	89
228	-L30603643	LIB3060-046-Q1-K1-G7	LIB3060	g169037	BLASTX	155	1e-44	66
229	1182	LIB3079-006-Q1-K1-H8	LIB3079	g22632	BLASTN	598	1e-39	65
230	28633	LIB3062-015-Q1-K1-G12	LIB3062	g1208898	BLASTX	116	1e-24	45
231	38	LIB3061-025-Q1-K1-C9	LIB3061	g22144	BLASTN	895	1e-133	94
232	38	LIB3059-020-Q1-K1-H3	LIB3059	g22144	BLASTN	745	1e-53	98

# MAIZE FRUCTOSE-1,6-BISPHOSPHATASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
233	-700262935	700262935H1	SATMON017	g3041775	BLASTX	184	1e-18	94
234	-700432173	700432173H1	SATMONN01	g1790679	BLASTX	123	1e-16	56
235	-700455709	700455709H1	SATMON029	g3041776	BLASTN	597	1e-40	85
236	-700573083	700573083H1	SATMON030	g3041775	BLASTX	69	1e-10	64
237	12846	700101851H1	SATMON009	g3041776	BLASTN	1312	1e-100	91
238	12846	700101541H1	SATMON009	g3041776	BLASTN	1252	1e-95	90
239	12846	700581510H1	SATMON031	g3041776	BLASTN	872	1e-82	90
240	15627	700046054H1	SATMON004	g21736	BLASTN	1213	1e-92	91
241	15627	700421605H1	SATMONN01	g3041776	BLASTN	664	1e-77	90
242	15627	700445495H1	SATMON027	g21736	BLASTN	1004	1e-74	84
243	15627	700042188H1	SATMON004	g3041776	BLASTN	875	1e-64	88
244	16870	700100752H1	SATMON009	g3041776	BLASTN	257	1e-33	75
245	16870	700044805H1	SATMON004	g3041776	BLASTN	194	1e-14	76
246	16870	700099217H1	SATMON009	g21736	BLASTN	246	1e-9	59
247	5480	700442189H1	SATMON026	g3041774	BLASTN	536	1e-54	93
248	8243	700264654H1	SATMON017	g3041774	BLASTN	942	1e-69	84
249	8243	700479624H1	SATMON034	g3041774	BLASTN	902	1e-66	82
250	8243	700448974H1	SATMON028	g3041774	BLASTN	876	1e-64	84
251	-L1485381	LIB148-057-Q1-E1-E6	LIB148	g440591	BLASTX	80	1e-30	63
252	-L30662839	LIB3066-035-Q1-K1-F11	LIB3066	g3041774	BLASTN	215	1e-15	77
253	-L362913	LIB36-013-Q1-E1-D10	LIB36	g3041776	BLASTN	937	1e-69	88
254	-L832444	LIB83-005-Q1-E1-D2	LIB83	g3041776	BLASTN	575	1e-37	93
255	12846	LIB83-008-Q1-E1-A8	LIB83	g3041776	BLASTN	1610	1e-135	92
256	12846	LIB3078-003-Q1-K1-C7	LIB3078	g3041776	BLASTN	873	1e-98	93
257	16870	LIB3060-052-Q1-K1-D11	LIB3060	g21736	BLASTN	377	1e-66	70
258	26002	LIB83-008-	LIB83	g3041776	BLASTN	378	1e-20	86

## Q1-E1-B10

## MAIZE FRUCTOSE-6-PHOSPHATE,2-KINASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
259	-700093724	700093724H1	SATMON008	g3170230	BLASTX	123	1e-21	53
260	-700099547	700099547H1	SATMON009	g3309582	BLASTN	630	1e-43	80
261	-700100682	700100682H1	SATMON009	g3170230	BLASTX	269	1e-39	65
262	-700173085	700173085H1	SATMON013	g2286154	BLASTN	1165	1e-88	100
263	-700217623	700217623H1	SATMON016	g3170229	BLASTN	593	1e-40	73
264	-700219340	700219340H1	SATMON011	g3170230	BLASTX	190	1e-20	56
265	-700265353	700265353H1	SATMON017	g2286154	BLASTN	1268	1e-107	98
266	-700379777	700379777H1	SATMON021	g3309582	BLASTN	905	1e-66	76
267	-700620963	700620963H1	SATMON034	g2286154	BLASTN	376	1e-52	85
268	-701159590	701159590H1	SATMONN04	g3309582	BLASTN	682	1e-48	73
269	20094	700209789H1	SATMON016	g2286154	BLASTN	1093	1e-96	92
270	20094	700550375H1	SATMON022	g3309582	BLASTN	780	1e-58	81
271	29193	700021150H1	SATMON001	g2286154	BLASTN	466	1e-75	92
272	-L30593297	LIB3059-029-Q1-K1-B3	LIB3059	g2286154	BLASTN	401	1e-22	70
273	-L30614892	LIB3061-021-Q1-K1-G9	LIB3061	g2286154	BLASTN	469	1e-38	79
274	-L30623700	LIB3062-031-Q1-K1-E8	LIB3062	g3170229	BLASTN	230	1e-10	70
275	29193	LIB83-007-Q1-E1-C11	LIB83	g2286154	BLASTN	595	1e-113	90

## MAIZE PHOSPHOGLUCOISOMERASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
276	-700086021	700086021H1	SATMON011	g1100771	BLASTX	225	1e-28	51
277	-700169489	700169489H1	SATMON013	g1100771	BLASTX	152	1e-13	59
278	-700222638	700222638H1	SATMON011	g1100771	BLASTX	256	1e-28	60
279	-700445574	700445574H1	SATMON027	g1100771	BLASTX	143	1e-12	54
280	-700475232	700475232H1	SATMON025	g596022	BLASTN	845	1e-61	90
281	-700612774	700612774H1	SATMON033	g596022	BLASTN	1574	1e-122	95
282	14393	700222547H1	SATMON011	g1100771	BLASTX	239	1e-25	60
283	14393	700220357H1	SATMON011	g1100771	BLASTX	218	1e-23	68
284	14393	700050317H1	SATMON003	g1100771	BLASTX	120	1e-22	63
285	14393	700163544H1	SATMON013	g1100771	BLASTX	214	1e-22	62
286	15724	700207164H1	SATMON017	g1100771	BLASTX	135	1e-17	67
287	15724	700552402H1	SATMON022	g1100771	BLASTX	135	1e-11	60
288	15724	700086085H1	SATMON011	g1100771	BLASTX	137	1e-11	45
289	20643	700577051H1	SATMON031	g1100771	BLASTX	241	1e-26	66
290	20643	700201592H1	SATMON003	g1100771	BLASTX	113	1e-19	45
291	20643	700576644H1	SATMON030	g1100771	BLASTX	113	1e-17	43
292	2351	700208928H1	SATMON016	g1100771	BLASTX	274	1e-43	73
293	2351	700240758H1	SATMON010	g1100771	BLASTX	283	1e-43	79
294	2351	700352502H1	SATMON023	g1100771	BLASTX	197	1e-36	70
295	2351	700581930H1	SATMON031	g1100771	BLASTX	164	1e-34	72
296	2351	700028642H1	SATMON003	g1100771	BLASTX	294	1e-33	65
297	2351	700106092H1	SATMON010	g1100771	BLASTX	294	1e-33	62
298	2351	700082102H1	SATMON011	g1100771	BLASTX	300	1e-33	62

299	2351	700083446H1	SATMON011	g1100771	BLASTX	274	1e-30	65
300	2351	700580585H1	SATMON031	g1100771	BLASTX	163	1e-29	69
301	2351	700550608H1	SATMON022	g1100771	BLASTX	265	1e-29	61
302	2351	700106079H1	SATMON010	g1100771	BLASTX	261	1e-28	54
303	2351	700244248H1	SATMON010	g1100771	BLASTX	238	1e-25	67
304	2351	700152233H1	SATMON007	g1100771	BLASTX	167	1e-22	72
305	2351	700455043H1	SATMON029	g1100771	BLASTX	168	1e-21	68
306	2351	700615809H1	SATMON033	g1100771	BLASTX	207	1e-21	66
307	2351	701165320H1	SATMONN04	g1100771	BLASTX	122	1e-14	63
308	32930	700042996H1	SATMON004	g596022	BLASTN	476	1e-95	98
309	4222	700222539H1	SATMON011	g596022	BLASTN	1160	1e-87	100
310	4222	700104023H1	SATMON010	g596022	BLASTN	1060	1e-84	100
311	4222	700101580H1	SATMON009	g596022	BLASTN	871	1e-74	99
312	4222	700473395H1	SATMON025	g596022	BLASTN	368	1e-46	95
313	4222	700800179H1	SATMON036	g596022	BLASTN	240	1e-11	100
314	8858	700221523H1	SATMON011	g1100771	BLASTX	278	1e-31	59
315	895	700100965H1	SATMON009	g596022	BLASTN	1611	1e-125	99
316	895	700620985H1	SATMON034	g596022	BLASTN	1418	1e-114	98
317	895	700082062H1	SATMON011	g596022	BLASTN	1365	1e-110	97
318	895	700573782H1	SATMON030	g596022	BLASTN	920	1e-107	98
319	895	700236138H1	SATMON010	g596022	BLASTN	1395	1e-107	100
320	895	700086336H1	SATMON011	g596022	BLASTN	1370	1e-105	100
321	895	700801467H1	SATMON036	g596022	BLASTN	1249	1e-99	95
322	895	700801458H1	SATMON036	g596022	BLASTN	1245	1e-98	100
323	895	700475024H1	SATMON025	g596022	BLASTN	1162	1e-97	93
324	895	700243164H1	SATMON010	g596022	BLASTN	1105	1e-96	100
325	895	700804665H1	SATMON036	g596022	BLASTN	1266	1e-96	99
326	895	700021931H1	SATMON001	g596022	BLASTN	1126	1e-84	99
327	895	700805540H1	SATMON036	g596022	BLASTN	776	1e-55	99
328	895	700172576H1	SATMON013	g596022	BLASTN	571	1e-38	98
329	895	700105116H1	SATMON010	g596022	BLASTN	558	1e-37	99
330	895	700472931H1	SATMON025	g596022	BLASTN	379	1e-31	97
331	20643	LIB3069-009-Q1-K1-B3	LIB3069	g1100771	BLASTX	215	1e-44	50
332	2351	LIB3079-007-Q1-K1-C11	LIB3079	g1100771	BLASTX	304	1e-77	72
333	32930	LIB189-001-Q1-E1-E4	LIB189	g596022	BLASTN	794	1e-115	95
334	4222	LIB3079-001-Q1-K1-H7	LIB3079	g596022	BLASTN	1132	1e-101	89
335	895	LIB148-049-Q1-E1-D6	LIB148	g596022	BLASTN	2194	1e-178	97
336	895	LIB3066-052-Q1-K1-G8	LIB3066	g596022	BLASTN	2178	1e-172	97
337	895	LIB148-016-Q1-E1-G5	LIB148	g596022	BLASTN	1567	1e-161	99
338	895	LIB143-032-Q1-E1-E10	LIB143	g596022	BLASTN	1914	1e-155	99
339	895	LIB3061-013-Q1-K1-F7	LIB3061	g596022	BLASTN	1738	1e-136	88
340	895	LIB143-047-Q1-E1-D4	LIB143	g596022	BLASTN	1490	1e-119	88

**MAIZE VACUOLAR H<sup>+</sup>-TRANSLOCATING-PYROPHOSPHATASE**

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
341	-700163331	700163331H1	SATMON013	g534915	BLASTN	751	1e-53	77
342	-700171438	700171438H1	SATMON013	g2258073	BLASTN	256	1e-10	76
343	-700202576	700202576H1	SATMON003	g2668746	BLASTX	214	1e-23	84
344	-700206487	700206487H1	SATMON003	g2570501	BLASTX	174	1e-17	86
345	-700217292	700217292H1	SATMON016	g2668746	BLASTX	214	1e-23	100
346	-700240889	700240889H1	SATMON010	g2570500	BLASTN	639	1e-47	84
347	-700347658	700347658H1	SATMON023	g2668746	BLASTX	215	1e-23	95
348	-700454151	700454151H1	SATMON029	g2668745	BLASTN	172	1e-10	90
349	-700454532	700454532H1	SATMON029	g2668745	BLASTN	259	1e-38	93
350	-700552133	700552133H1	SATMON022	g457744	BLASTX	176	1e-19	68
351	-700611864	700611864H1	SATMON022	g2668745	BLASTN	203	1e-9	84
352	107	700622451H1	SATMON034	g2668745	BLASTN	1645	1e-129	100
353	107	700571235H1	SATMON030	g2668745	BLASTN	1406	1e-125	98
354	107	700266126H1	SATMON017	g2668745	BLASTN	1145	1e-121	100
355	107	700621607H1	SATMON034	g2668745	BLASTN	1375	1e-121	99
356	107	700345080H1	SATMON021	g2668745	BLASTN	1195	1e-117	100
357	107	700624257H1	SATMON034	g2668745	BLASTN	825	1e-115	100
358	107	700030359H1	SATMON003	g2668745	BLASTN	1470	1e-114	100
359	107	700214462H1	SATMON016	g2668745	BLASTN	1223	1e-110	98
360	107	700356050H1	SATMON024	g2668745	BLASTN	1430	1e-110	100
361	107	701181128H1	SATMONN06	g2668745	BLASTN	1368	1e-105	98
362	107	700349795H1	SATMON023	g2668745	BLASTN	1370	1e-105	95
363	107	700473278H1	SATMON025	g2668745	BLASTN	1355	1e-104	100
364	107	700157057H1	SATMON012	g2668745	BLASTN	1345	1e-103	100
365	107	700622505H1	SATMON034	g2668745	BLASTN	762	1e-100	96
366	107	700219661H1	SATMON011	g2668745	BLASTN	942	1e-98	99
367	107	700619032H1	SATMON034	g2668745	BLASTN	989	1e-98	96
368	107	700620065H1	SATMON034	g2668745	BLASTN	1069	1e-98	94
369	107	700569179H1	SATMON030	g2668745	BLASTN	1233	1e-97	98
370	107	700156773H1	SATMON012	g2668745	BLASTN	1276	1e-97	99
371	107	700207120H1	SATMON017	g2668745	BLASTN	740	1e-96	99
372	107	700030407H1	SATMON003	g2668745	BLASTN	480	1e-95	98
373	107	700457309H1	SATMON029	g2668745	BLASTN	979	1e-95	99
374	107	700195681H1	SATMON014	g2668745	BLASTN	1246	1e-95	99
375	107	700444838H1	SATMON027	g2668745	BLASTN	1249	1e-95	96
376	107	700581619H1	SATMON031	g2668745	BLASTN	943	1e-94	96
377	107	700351021H1	SATMON023	g2668745	BLASTN	853	1e-91	92
378	107	700205723H1	SATMON003	g2668745	BLASTN	1138	1e-91	95
379	107	700159712H1	SATMON012	g2668745	BLASTN	1199	1e-91	94
380	107	700158937H1	SATMON012	g2668745	BLASTN	1132	1e-90	96
381	107	700336255H1	SATMON019	g2668745	BLASTN	489	1e-85	94
382	107	700422922H1	SATMONN01	g2668745	BLASTN	642	1e-84	95
383	107	700347429H1	SATMON023	g2668745	BLASTN	891	1e-83	92
384	107	700350695H1	SATMON023	g2668745	BLASTN	960	1e-83	91
385	107	700212988H1	SATMON016	g2668745	BLASTN	988	1e-82	96
386	107	700345278H1	SATMON021	g2668745	BLASTN	989	1e-82	95
387	107	700264475H1	SATMON017	g2668745	BLASTN	1089	1e-82	99
388	107	700211923H1	SATMON016	g2668745	BLASTN	991	1e-81	94
389	107	700620974H1	SATMON034	g2668745	BLASTN	907	1e-80	92
390	107	700156401H1	SATMON012	g2668745	BLASTN	1058	1e-79	90
391	107	700172547H1	SATMON013	g2668745	BLASTN	1042	1e-78	96
392	107	700552384H1	SATMON022	g2668745	BLASTN	916	1e-76	96



393	107	700219926H1	SATMON011	g2668745	BLASTN	1005	1e-75	100
394	107	700357492H1	SATMON024	g2668745	BLASTN	610	1e-74	99
395	107	700343365H1	SATMON021	g2668745	BLASTN	891	1e-74	94
396	107	700018618H1	SATMON001	g2668745	BLASTN	1001	1e-74	93
397	107	700570755H1	SATMON030	g2668745	BLASTN	845	1e-71	93
398	107	700194777H1	SATMON014	g2668745	BLASTN	940	1e-69	100
399	107	700453790H1	SATMON029	g2668745	BLASTN	925	1e-68	92
400	107	700197306H1	SATMON014	g2668745	BLASTN	928	1e-68	85
401	107	700355750H1	SATMON024	g2668745	BLASTN	393	1e-66	93
402	107	700172940H1	SATMON013	g2668745	BLASTN	902	1e-66	97
403	107	700102133H1	SATMON010	g2668745	BLASTN	850	1e-62	100
404	107	700350332H1	SATMON023	g2668745	BLASTN	539	1e-57	97
405	107	700450285H1	SATMON028	g2668745	BLASTN	750	1e-53	100
406	107	700165003H1	SATMON013	g2668745	BLASTN	548	1e-52	83
407	107	700016136H1	SATMON001	g2668745	BLASTN	527	1e-50	85
408	107	700171557H1	SATMON013	g2668745	BLASTN	714	1e-50	95
409	107	700238156H1	SATMON010	g2668745	BLASTN	715	1e-50	96
410	107	700425175H1	SATMONN01	g2668745	BLASTN	698	1e-49	94
411	107	700354402H1	SATMON024	g2668745	BLASTN	616	1e-48	91
412	107	700159204H1	SATMON012	g2668745	BLASTN	617	1e-42	94
413	107	700623602H1	SATMON034	g2668745	BLASTN	460	1e-38	100
414	107	700612844H1	SATMON033	g2668745	BLASTN	421	1e-36	84
415	107	700621062H2	SATMON034	g2668745	BLASTN	285	1e-25	89
416	107	700335685H1	SATMON019	g2668745	BLASTN	339	1e-25	91
417	13843	700334949H1	SATMON019	g2570500	BLASTN	680	1e-55	83
418	13843	700346817H1	SATMON021	g2570500	BLASTN	705	1e-54	83
419	13843	700103380H1	SATMON010	g2570500	BLASTN	710	1e-54	83
420	13843	700348280H1	SATMON023	g2570500	BLASTN	669	1e-51	83
421	13843	700453203H1	SATMON028	g2570500	BLASTN	659	1e-50	82
422	13843	700381101H1	SATMON023	g2570500	BLASTN	621	1e-47	82
423	13843	700347617H1	SATMON023	g2570500	BLASTN	592	1e-44	85
424	13843	700043259H1	SATMON004	g2570500	BLASTN	530	1e-39	84
425	13843	701184447H1	SATMONN06	g2570500	BLASTN	481	1e-35	78
426	21076	700241354H1	SATMON010	g166634	BLASTX	201	1e-20	58
427	24066	700423113H1	SATMONN01	g457744	BLASTX	124	1e-23	54
428	24266	700577157H1	SATMON031	g2570500	BLASTN	1001	1e-74	89
429	2531	700099364H1	SATMON009	g2570500	BLASTN	669	1e-51	86
430	2531	700336387H1	SATMON019	g2570500	BLASTN	389	1e-47	85
431	2531	700217095H1	SATMON016	g2570500	BLASTN	451	1e-33	86
432	2531	700155869H1	SATMON007	g2570500	BLASTN	385	1e-27	89
433	2531	700575534H1	SATMON030	g2570500	BLASTN	365	1e-26	88
434	2531	700163562H1	SATMON013	g2570501	BLASTX	145	1e-24	94
435	32364	700204306H1	SATMON003	g2668745	BLASTN	471	1e-28	74
436	32856	700166756H1	SATMON013	g534915	BLASTN	744	1e-53	76
437	32856	700042535H1	SATMON004	g534915	BLASTN	644	1e-44	73
438	3384	700237775H1	SATMON010	g2258073	BLASTN	911	1e-67	81
439	3384	700342456H1	SATMON021	g2258073	BLASTN	648	1e-64	78
440	3384	700073654H1	SATMON007	g2668745	BLASTN	860	1e-63	78
441	3384	700577805H1	SATMON031	g2258073	BLASTN	840	1e-61	78
442	3384	700028881H1	SATMON003	g534915	BLASTN	835	1e-60	78
443	3384	700215076H1	SATMON016	g534915	BLASTN	824	1e-59	78
444	3384	700017479H1	SATMON001	g534915	BLASTN	766	1e-55	80
445	3384	700204495H1	SATMON003	g534915	BLASTN	373	1e-51	81
446	3384	700206347H1	SATMON003	g2706449	BLASTN	685	1e-48	80

447	3384	700351040H1	SATMON023	g2706449	BLASTN	436	1e-45	78
448	3384	700345264H1	SATMON021	g2706449	BLASTN	616	1e-42	82
449	3384	700196795H1	SATMON014	g2570500	BLASTN	579	1e-39	80
450	3384	700019241H1	SATMON001	g2706449	BLASTN	583	1e-39	78
451	3384	700018612H1	SATMON001	g2668745	BLASTN	518	1e-34	76
452	3384	700102142H1	SATMON010	g2668745	BLASTN	539	1e-34	78
453	3384	700348430H1	SATMON023	g534915	BLASTN	489	1e-30	78
454	3384	700337745H1	SATMON020	g2706449	BLASTN	471	1e-28	79
455	3384	700439515H1	SATMON026	g534915	BLASTN	437	1e-27	75
456	3384	700074977H1	SATMON007	g534915	BLASTN	434	1e-25	76
457	3384	700615213H1	SATMON033	g2570501	BLASTX	125	1e-21	93
458	3384	700074109H1	SATMON007	g2668746	BLASTX	197	1e-20	72
459	3384	700549517H1	SATMON022	g2668746	BLASTX	172	1e-17	75
460	3384	700030347H1	SATMON003	g2668746	BLASTX	171	1e-16	77
461	3384	700221176H1	SATMON011	g2668746	BLASTX	171	1e-16	77
462	3384	700433360H1	SATMONN01	g2668746	BLASTX	95	1e-13	74
463	5000	700026151H1	SATMON003	g2903	BLASTX	261	1e-28	54
464	5000	700347165H1	SATMON021	g2624379	BLASTX	223	1e-24	51
465	5000	700430341H1	SATMONN01	g2903	BLASTX	185	1e-18	56
466	5000	700457781H1	SATMON029	g2903	BLASTX	133	1e-16	49
467	5861	700104993H1	SATMON010	g2258073	BLASTN	456	1e-27	73
468	5861	700203452H1	SATMON003	g2258073	BLASTN	428	1e-26	72
469	-L1431590	LIB143-006-Q1-E1-C9	LIB143	g16347	BLASTN	286	1e-13	61
470	-L1433414	LIB143-026-Q1-E1-C3	LIB143	g2258073	BLASTN	480	1e-29	70
471	-L1482832	LIB148-009-Q1-E1-D8	LIB148	g2258073	BLASTN	1086	1e-81	78
472	-L30674379	LIB3067-042-Q1-K1-H8	LIB3067	g2668745	BLASTN	305	1e-21	68
473	-L30675678	LIB3067-034-Q1-K1-E3	LIB3067	g2706449	BLASTN	286	1e-12	73
474	107	LIB3059-036-Q1-K1-B10	LIB3059	g2668745	BLASTN	1965	1e-166	100
475	107	LIB3061-035-Q1-K1-C9	LIB3061	g2668745	BLASTN	948	1e-138	93
476	107	LIB3061-032-Q1-K1-A12	LIB3061	g2668745	BLASTN	1685	1e-138	96
477	107	LIB3062-044-Q1-K1-F8	LIB3062	g2668745	BLASTN	1492	1e-134	95
478	107	LIB3068-025-Q1-K1-E5	LIB3068	g2668745	BLASTN	1687	1e-132	96
479	107	LIB3067-022-Q1-K1-D11	LIB3067	g2668745	BLASTN	1581	1e-128	91
480	107	LIB3067-016-Q1-K1-G4	LIB3067	g2668745	BLASTN	1305	1e-126	97
481	107	LIB3067-029-Q1-K1-C6	LIB3067	g2668745	BLASTN	1560	1e-125	90
482	107	LIB189-031-Q1-E1-D3	LIB189	g2668745	BLASTN	897	1e-81	85
483	24066	LIB3069-047-Q1-K1-C4	LIB3069	g166634	BLASTX	173	1e-45	55
484	24266	LIB3069-006-Q1-K1-F4	LIB3069	g2570500	BLASTN	717	1e-57	83

485	293	LIB3068-043-Q1-K1-A2	LIB3068	g633598	BLASTN	552	1e-34	78
486	32364	LIB3066-001-Q1-K1-B7	LIB3066	g2668745	BLASTN	612	1e-40	73
487	32856	LIB189-028-Q1-E1-C4	LIB189	g534915	BLASTN	986	1e-73	73
488	3384	LIB143-026-Q1-E1-C1	LIB143	g534915	BLASTN	1284	1e-98	78
489	3384	LIB3068-013-Q1-K1-H2	LIB3068	g534915	BLASTN	1074	1e-80	78
490	3384	LIB3062-033-Q1-K1-D2	LIB3062	g2668745	BLASTN	1009	1e-75	76
491	3384	LIB83-002-Q1-E1-D2	LIB83	g2706449	BLASTN	820	1e-59	78
492	3384	LIB3062-057-Q1-K1-B7	LIB3062	g2668745	BLASTN	801	1e-58	73
493	3384	LIB3062-001-Q1-K2-H5	LIB3062	g16347	BLASTN	802	1e-57	77
494	3384	LIB189-022-Q1-E1-D5	LIB189	g2668745	BLASTN	646	1e-43	75
495	3384	LIB189-012-Q1-E1-F4	LIB189	g2570501	BLASTX	138	1e-32	72
496	5000	LIB36-015-Q1-E1-D6	LIB36	g2624379	BLASTX	236	1e-41	51
497	5000	LIB83-016-Q1-E1-H7	LIB83	g4198	BLASTN	534	1e-33	61

#### MAIZE PYROPHOSPHATE-DEPENDENT FRUCTOSE-6-PHOSPHATE PHOSPHOTRANSFERASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
498	-700208959	700208959H1	SATMON016	g169538	BLASTX	107	1e-19	50
499	-700237606	700237606H1	SATMON010	g169538	BLASTX	114	1e-11	62
500	3456	700083478H1	SATMON011	g169538	BLASTX	121	1e-39	88
501	3652	700242182H1	SATMON010	g169538	BLASTX	155	1e-13	82
502	4965	700475352H1	SATMON025	g169538	BLASTX	123	1e-9	69
503	4965	700550752H1	SATMON022	g169538	BLASTX	123	1e-9	69
504	5359	700347441H1	SATMON023	g169538	BLASTX	139	1e-11	70
505	-L30594734	LIB3059-018-Q1-K1-H3	LIB3059	g169538	BLASTX	145	1e-49	83
506	-L30622375	LIB3062-009-Q1-K1-B3	LIB3062	g169538	BLASTX	157	1e-30	65
507	32156	LIB189-021-Q1-E1-G8	LIB189	g169538	BLASTX	123	1e-25	78

#### MAIZE INVERTASES

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
508	-700240132	700240132H1	SATMON010	g397631	BLASTX	134	1e-11	74
509	1923	700574932H1	SATMON030	g393390	BLASTX	152	1e-14	65
510	4355	700379641H1	SATMON021	g1177601	BLASTX	175	1e-19	85

#### MAIZE SUCROSE SYNTHASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	%Ident
511	-700151470	700151470H1	SATMON007	g1196837	BLASTX	197	1e-27	64
512	-700214035	700214035H1	SATMON016	g22485	BLASTN	523	1e-34	79
513	-700262270	700262270H1	SATMON017	g2570066	BLASTN	866	1e-63	76
514	-700334686	700334686H1	SATMON019	g1100216	BLASTN	424	1e-31	88
515	-700381593	700381593H1	SATMON023	g22485	BLASTN	219	1e-13	97
516	-700404808	700404808H1	SATMON026	g2570066	BLASTN	859	1e-70	82
517	-700456905	700456905H1	SATMON029	g22485	BLASTN	528	1e-64	90
518	-700571529	700571529H1	SATMON030	g19106	BLASTX	139	1e-24	56
519	-700576567	700576567H1	SATMON030	g22485	BLASTN	285	1e-14	92
520	-700800659	700800659H1	SATMON036	g22485	BLASTN	558	1e-37	97
521	-700802941	700802941H1	SATMON036	g22485	BLASTN	316	1e-29	97
522	-701181030	701181030H1	SATMONN06	g2606080	BLASTN	669	1e-46	72
523	13723	700203023H1	SATMON003	g2570066	BLASTN	820	1e-68	84
524	13723	700215119H1	SATMON016	g2570066	BLASTN	680	1e-47	86
525	13723	700473266H1	SATMON025	g2570066	BLASTN	537	1e-35	85
526	15661	700440404H1	SATMON026	g2570066	BLASTN	364	1e-36	74
527	15661	700168252H1	SATMON013	g16525	BLASTN	433	1e-27	80
528	20925	700551647H1	SATMON022	g2570066	BLASTN	307	1e-35	73
529	20925	700257052H1	SATMON017	g2570067	BLASTX	118	1e-9	64
530	20934	700217752H1	SATMON016	g514945	BLASTN	1397	1e-107	98
531	20934	700332156H1	SATMON019	g514945	BLASTN	589	1e-97	95
532	30444	700257522H1	SATMON017	g1100216	BLASTN	760	1e-54	95
533	32909	700264718H1	SATMON017	g2570066	BLASTN	702	1e-57	76
534	405	700091402H1	SATMON011	g514945	BLASTN	1830	1e-143	100
535	405	700572549H1	SATMON030	g514945	BLASTN	1658	1e-129	99
536	405	700203058H1	SATMON003	g22485	BLASTN	1360	1e-127	100
537	405	700091753H1	SATMON011	g514945	BLASTN	1245	1e-126	99
538	405	700090929H1	SATMON011	g514945	BLASTN	1620	1e-126	100
539	405	700091711H1	SATMON011	g514945	BLASTN	1621	1e-126	99
540	405	700084254H1	SATMON011	g514945	BLASTN	1600	1e-124	100
541	405	700082305H1	SATMON011	g514945	BLASTN	1601	1e-124	99
542	405	700048236H1	SATMON003	g22485	BLASTN	1583	1e-123	99
543	405	700086713H1	SATMON011	g514945	BLASTN	1584	1e-123	99
544	405	700049353H1	SATMON003	g514945	BLASTN	1586	1e-123	99
545	405	700082766H1	SATMON011	g22485	BLASTN	1589	1e-123	98
546	405	700086055H1	SATMON011	g514945	BLASTN	1590	1e-123	100
547	405	700215105H1	SATMON016	g514945	BLASTN	1590	1e-123	100
548	405	700104149H1	SATMON010	g22485	BLASTN	1594	1e-123	98
549	405	700101601H1	SATMON009	g514945	BLASTN	1270	1e-122	100
550	405	700206869H1	SATMON003	g22485	BLASTN	1574	1e-122	97
551	405	700088163H1	SATMON011	g22485	BLASTN	1581	1e-122	99
552	405	700089166H1	SATMON011	g514945	BLASTN	1565	1e-121	100
553	405	700266251H1	SATMON017	g514945	BLASTN	1570	1e-121	100
554	405	700332710H1	SATMON019	g514945	BLASTN	1570	1e-121	100
555	405	700571106H1	SATMON030	g514945	BLASTN	1227	1e-120	98
556	405	700081893H1	SATMON011	g514945	BLASTN	1550	1e-120	98
557	405	700074739H1	SATMON007	g514945	BLASTN	1550	1e-120	100
558	405	700095163H1	SATMON008	g514945	BLASTN	1555	1e-120	100
559	405	700612766H1	SATMON033	g514945	BLASTN	883	1e-119	96
560	405	700267271H1	SATMON017	g514945	BLASTN	1535	1e-119	100
561	405	700083175H1	SATMON011	g514945	BLASTN	1535	1e-119	100
562	405	700088993H1	SATMON011	g22485	BLASTN	1545	1e-119	98

563	405	700094087H1	SATMON008	g22485	BLASTN	1526	1e-118	99
564	405	700086708H1	SATMON011	g514945	BLASTN	1529	1e-118	97
565	405	700090671H1	SATMON011	g514945	BLASTN	1530	1e-118	100
566	405	700209809H1	SATMON016	g22485	BLASTN	1532	1e-118	99
567	405	700084625H1	SATMON011	g514945	BLASTN	1533	1e-118	99
568	405	700089718H1	SATMON011	g514945	BLASTN	1120	1e-117	100
569	405	700213014H1	SATMON016	g514945	BLASTN	1405	1e-117	100
570	405	700086555H1	SATMON011	g514945	BLASTN	1514	1e-117	98
571	405	700475892H1	SATMON025	g514945	BLASTN	1516	1e-117	99
572	405	700047374H1	SATMON003	g22485	BLASTN	1516	1e-117	99
573	405	700090018H1	SATMON011	g514945	BLASTN	1519	1e-117	99
574	405	700076107H1	SATMON007	g514945	BLASTN	1520	1e-117	93
575	405	700213105H1	SATMON016	g514945	BLASTN	972	1e-116	99
576	405	700103806H1	SATMON010	g514945	BLASTN	1503	1e-116	99
577	405	700090748H1	SATMON011	g514945	BLASTN	1505	1e-116	100
578	405	700052006H1	SATMON003	g514945	BLASTN	1506	1e-116	99
579	405	700614963H1	SATMON033	g514945	BLASTN	957	1e-115	93
580	405	700337255H1	SATMON020	g22485	BLASTN	995	1e-115	97
581	405	700102778H1	SATMON010	g22485	BLASTN	1493	1e-115	99
582	405	700405466H1	SATMON029	g22485	BLASTN	1493	1e-115	99
583	405	700209634H1	SATMON016	g514945	BLASTN	1495	1e-115	100
584	405	700220467H1	SATMON011	g514945	BLASTN	1495	1e-115	100
585	405	700266637H1	SATMON017	g514945	BLASTN	1480	1e-114	100
586	405	700267579H1	SATMON017	g514945	BLASTN	1484	1e-114	99
587	405	700088475H1	SATMON011	g514945	BLASTN	1465	1e-113	100
588	405	700332618H1	SATMON019	g514945	BLASTN	1466	1e-113	99
589	405	700211347H1	SATMON016	g514945	BLASTN	1470	1e-113	100
590	405	700477206H1	SATMON025	g514945	BLASTN	1471	1e-113	99
591	405	700336768H1	SATMON019	g514945	BLASTN	1473	1e-113	99
592	405	700105305H1	SATMON010	g22485	BLASTN	1473	1e-113	99
593	405	700087114H1	SATMON011	g514945	BLASTN	1473	1e-113	99
594	405	700105366H1	SATMON010	g22485	BLASTN	1474	1e-113	98
595	405	700104831H1	SATMON010	g22485	BLASTN	825	1e-112	98
596	405	700620134H1	SATMON034	g22485	BLASTN	1179	1e-112	92
597	405	700211934H1	SATMON016	g22485	BLASTN	1215	1e-112	98
598	405	700096103H1	SATMON008	g514945	BLASTN	1391	1e-112	99
599	405	700264979H1	SATMON017	g514945	BLASTN	1454	1e-112	98
600	405	700053864H1	SATMON011	g514945	BLASTN	1455	1e-112	100
601	405	700211782H1	SATMON016	g514945	BLASTN	1460	1e-112	100
602	405	700102063H1	SATMON010	g22485	BLASTN	1461	1e-112	99
603	405	700207024H1	SATMON003	g514945	BLASTN	825	1e-111	100
604	405	700207970H1	SATMON016	g514945	BLASTN	1186	1e-111	98
605	405	700336624H1	SATMON019	g514945	BLASTN	1440	1e-111	100
606	405	700104357H1	SATMON010	g514945	BLASTN	1448	1e-111	98
607	405	700222053H1	SATMON011	g514945	BLASTN	1449	1e-111	99
608	405	700350806H1	SATMON023	g514945	BLASTN	660	1e-110	99
609	405	700091159H1	SATMON011	g514945	BLASTN	870	1e-110	100
610	405	700081810H1	SATMON011	g514945	BLASTN	926	1e-110	99
611	405	700102954H1	SATMON010	g514945	BLASTN	926	1e-110	97
612	405	700085307H1	SATMON011	g514945	BLASTN	1035	1e-110	100
613	405	700094295H1	SATMON008	g22485	BLASTN	1137	1e-110	96
614	405	700089176H1	SATMON011	g514945	BLASTN	1393	1e-110	97
615	405	700093643H1	SATMON008	g514945	BLASTN	1427	1e-110	95
616	405	700082421H1	SATMON011	g514945	BLASTN	1430	1e-110	98

617	405	700211788H1	SATMON016	g514945	BLASTN	1431	1e-110	99
618	405	700026724H1	SATMON003	g514945	BLASTN	1433	1e-110	97
619	405	700085275H1	SATMON011	g514945	BLASTN	1435	1e-110	100
620	405	700472161H1	SATMON025	g514945	BLASTN	755	1e-109	99
621	405	700084926H1	SATMON011	g514945	BLASTN	825	1e-109	100
622	405	700084592H1	SATMON011	g514945	BLASTN	920	1e-109	100
623	405	700053811H1	SATMON011	g514945	BLASTN	1296	1e-109	96
624	405	700216963H1	SATMON016	g514945	BLASTN	1415	1e-109	100
625	405	700085273H1	SATMON011	g22485	BLASTN	1416	1e-109	98
626	405	700082127H1	SATMON011	g514945	BLASTN	1420	1e-109	100
627	405	700085731H1	SATMON011	g514945	BLASTN	1425	1e-109	100
628	405	700088595H1	SATMON011	g22485	BLASTN	1426	1e-109	99
629	405	700470903H1	SATMON025	g514945	BLASTN	1426	1e-109	99
630	405	700265288H1	SATMON017	g514945	BLASTN	1375	1e-108	100
631	405	700072245H1	SATMON007	g514945	BLASTN	1404	1e-108	99
632	405	700347692H1	SATMON023	g514945	BLASTN	1405	1e-108	98
633	405	700214447H1	SATMON016	g514945	BLASTN	1406	1e-108	99
634	405	700476252H1	SATMON025	g514945	BLASTN	1407	1e-108	99
635	405	700336746H1	SATMON019	g514945	BLASTN	1409	1e-108	99
636	405	700053833H1	SATMON011	g514945	BLASTN	1410	1e-108	100
637	405	700094342H1	SATMON008	g514945	BLASTN	1410	1e-108	100
638	405	700202813H1	SATMON003	g514945	BLASTN	1032	1e-107	97
639	405	700050589H1	SATMON003	g514945	BLASTN	1035	1e-107	100
640	405	700050011H1	SATMON003	g514945	BLASTN	1078	1e-107	99
641	405	700215426H1	SATMON016	g514945	BLASTN	1189	1e-107	96
642	405	700472461H1	SATMON025	g514945	BLASTN	1392	1e-107	99
643	405	700336684H1	SATMON019	g22485	BLASTN	1393	1e-107	98
644	405	700449826H2	SATMON028	g514945	BLASTN	1395	1e-107	100
645	405	700216443H1	SATMON016	g514945	BLASTN	1396	1e-107	99
646	405	700240793H1	SATMON010	g514945	BLASTN	1399	1e-107	98
647	405	700215985H1	SATMON016	g514945	BLASTN	1400	1e-107	100
648	405	700336740H1	SATMON019	g514945	BLASTN	915	1e-106	99
649	405	700047958H1	SATMON003	g514945	BLASTN	987	1e-106	96
650	405	700085447H1	SATMON011	g514945	BLASTN	1030	1e-106	100
651	405	700084978H1	SATMON011	g514945	BLASTN	1121	1e-106	91
652	405	700800439H1	SATMON036	g22485	BLASTN	1379	1e-106	99
653	405	700219631H1	SATMON011	g514945	BLASTN	1380	1e-106	100
654	405	700220740H1	SATMON011	g514945	BLASTN	1380	1e-106	100
655	405	700243367H1	SATMON010	g514945	BLASTN	1381	1e-106	99
656	405	700220363H1	SATMON011	g514945	BLASTN	1387	1e-106	99
657	405	700215869H1	SATMON016	g514945	BLASTN	1390	1e-106	100
658	405	700216519H1	SATMON016	g514945	BLASTN	1131	1e-105	97
659	405	700052206H1	SATMON003	g514945	BLASTN	1264	1e-105	96
660	405	700094975H1	SATMON008	g514945	BLASTN	1368	1e-105	99
661	405	700220837H1	SATMON011	g514945	BLASTN	1369	1e-105	98
662	405	700221108H1	SATMON011	g514945	BLASTN	1370	1e-105	98
663	405	700222850H1	SATMON011	g514945	BLASTN	1370	1e-105	100
664	405	700214429H1	SATMON016	g514945	BLASTN	1373	1e-105	99
665	405	700473857H1	SATMON025	g514945	BLASTN	1375	1e-105	98
666	405	700213762H1	SATMON016	g514945	BLASTN	1378	1e-105	99
667	405	700405254H1	SATMON028	g22485	BLASTN	1242	1e-104	99
668	405	700029978H1	SATMON003	g22485	BLASTN	1324	1e-104	97
669	405	700238315H1	SATMON010	g514945	BLASTN	1355	1e-104	100
670	405	700241686H1	SATMON010	g514945	BLASTN	1358	1e-104	99

671	405	700237721H1	SATMON010	g22485	BLASTN	1360	1e-104	100
672	405	700217344H1	SATMON016	g514945	BLASTN	1360	1e-104	100
673	405	700030048H1	SATMON003	g514945	BLASTN	1363	1e-104	99
674	405	700211866H1	SATMON016	g514945	BLASTN	1363	1e-104	99
675	405	700214860H1	SATMON016	g514945	BLASTN	1365	1e-104	100
676	405	700085490H1	SATMON011	g514945	BLASTN	900	1e-103	98
677	405	700048568H1	SATMON003	g514945	BLASTN	980	1e-103	100
678	405	700381034H1	SATMON023	g22485	BLASTN	1269	1e-103	98
679	405	700220930H1	SATMON011	g514945	BLASTN	1347	1e-103	99
680	405	700030261H1	SATMON003	g514945	BLASTN	1353	1e-103	98
681	405	700081835H1	SATMON011	g22485	BLASTN	797	1e-102	98
682	405	700205270H1	SATMON003	g514945	BLASTN	1024	1e-102	94
683	405	700093612H1	SATMON008	g514945	BLASTN	1065	1e-102	99
684	405	700333392H1	SATMON019	g514945	BLASTN	1108	1e-102	97
685	405	700575385H1	SATMON030	g514945	BLASTN	1171	1e-102	96
686	405	700241061H1	SATMON010	g514945	BLASTN	1174	1e-102	99
687	405	700239916H1	SATMON010	g514945	BLASTN	1255	1e-102	100
688	405	700090248H1	SATMON011	g514945	BLASTN	1334	1e-102	98
689	405	700222923H1	SATMON011	g514945	BLASTN	1334	1e-102	98
690	405	700216993H1	SATMON016	g514945	BLASTN	1335	1e-102	100
691	405	700215984H1	SATMON016	g514945	BLASTN	1340	1e-102	100
692	405	700213182H1	SATMON016	g514945	BLASTN	1340	1e-102	98
693	405	700219845H1	SATMON011	g514945	BLASTN	1340	1e-102	100
694	405	700237762H1	SATMON010	g514945	BLASTN	1340	1e-102	100
695	405	700551043H1	SATMON022	g514945	BLASTN	1342	1e-102	99
696	405	700219254H1	SATMON011	g514945	BLASTN	1252	1e-101	99
697	405	700210348H1	SATMON016	g514945	BLASTN	1320	1e-101	97
698	405	700215089H1	SATMON016	g514945	BLASTN	1320	1e-101	100
699	405	700217251H1	SATMON016	g514945	BLASTN	1320	1e-101	100
700	405	700219240H1	SATMON011	g514945	BLASTN	1320	1e-101	100
701	405	700082094H1	SATMON011	g514945	BLASTN	1321	1e-101	99
702	405	700219385H1	SATMON011	g514945	BLASTN	1322	1e-101	99
703	405	700220052H1	SATMON011	g514945	BLASTN	1325	1e-101	100
704	405	700210366H1	SATMON016	g514945	BLASTN	1329	1e-101	93
705	405	700083089H1	SATMON011	g514945	BLASTN	1330	1e-101	98
706	405	700340286H1	SATMON020	g22485	BLASTN	677	1e-100	98
707	405	700221062H1	SATMON011	g514945	BLASTN	845	1e-100	98
708	405	700382272H1	SATMON024	g22485	BLASTN	958	1e-100	96
709	405	700209310H1	SATMON016	g514945	BLASTN	1187	1e-100	97
710	405	700052340H1	SATMON003	g514945	BLASTN	1188	1e-100	94
711	405	700467851H1	SATMON025	g22485	BLASTN	1245	1e-100	97
712	405	700088014H1	SATMON011	g514945	BLASTN	1270	1e-100	98
713	405	700214596H1	SATMON016	g514945	BLASTN	1295	1e-100	100
714	405	700157215H1	SATMON012	g22485	BLASTN	1310	1e-100	98
715	405	700223892H1	SATMON011	g514945	BLASTN	1310	1e-100	100
716	405	700218981H1	SATMON011	g514945	BLASTN	1310	1e-100	100
717	405	700081945H1	SATMON011	g514945	BLASTN	1311	1e-100	96
718	405	700217817H1	SATMON016	g514945	BLASTN	1315	1e-100	100
719	405	700469042H1	SATMON025	g514945	BLASTN	561	1e-99	98
720	405	700474709H1	SATMON025	g514945	BLASTN	801	1e-99	99
721	405	700201736H1	SATMON003	g514945	BLASTN	1168	1e-99	98
722	405	700223516H1	SATMON011	g514945	BLASTN	1201	1e-99	99
723	405	700453941H1	SATMON029	g22485	BLASTN	1295	1e-99	95
724	405	700212970H1	SATMON016	g514945	BLASTN	1297	1e-99	99

725	405	700215662H1	SATMON016	g22485	BLASTN	1297	1e-99	99
726	405	700802209H1	SATMON036	g22485	BLASTN	1300	1e-99	98
727	405	700343716H1	SATMON021	g514945	BLASTN	1300	1e-99	100
728	405	700223322H1	SATMON011	g514945	BLASTN	1300	1e-99	100
729	405	700217238H1	SATMON016	g514945	BLASTN	1300	1e-99	100
730	405	700195066H1	SATMON014	g22485	BLASTN	1300	1e-99	98
731	405	700072395H1	SATMON007	g514945	BLASTN	1301	1e-99	95
732	405	700212752H1	SATMON016	g22485	BLASTN	1305	1e-99	98
733	405	700222204H1	SATMON011	g514945	BLASTN	1305	1e-99	100
734	405	700550572H1	SATMON022	g22485	BLASTN	713	1e-98	97
735	405	700213879H1	SATMON016	g514945	BLASTN	866	1e-98	99
736	405	700551585H1	SATMON022	g514945	BLASTN	916	1e-98	99
737	405	700195025H1	SATMON014	g22485	BLASTN	1283	1e-98	98
738	405	700800710H1	SATMON036	g22485	BLASTN	1283	1e-98	98
739	405	700222985H1	SATMON011	g22485	BLASTN	1283	1e-98	98
740	405	700798823H1	SATMON036	g22485	BLASTN	1284	1e-98	98
741	405	700104391H1	SATMON010	g22485	BLASTN	1289	1e-98	98
742	405	700466592H1	SATMON025	g22485	BLASTN	1289	1e-98	95
743	405	700027037H1	SATMON003	g514945	BLASTN	919	1e-97	91
744	405	700214371H1	SATMON016	g514945	BLASTN	1033	1e-97	95
745	405	700799077H1	SATMON036	g22485	BLASTN	1091	1e-97	99
746	405	700467028H1	SATMON025	g514945	BLASTN	1103	1e-97	98
747	405	700219393H1	SATMON011	g514945	BLASTN	1250	1e-97	99
748	405	700197602H1	SATMON014	g22485	BLASTN	1273	1e-97	97
749	405	700801226H1	SATMON036	g22485	BLASTN	1276	1e-97	99
750	405	700216371H1	SATMON016	g514945	BLASTN	1278	1e-97	98
751	405	700805695H1	SATMON036	g22485	BLASTN	1280	1e-97	98
752	405	700334076H1	SATMON019	g514945	BLASTN	503	1e-96	98
753	405	700082647H1	SATMON011	g514945	BLASTN	735	1e-96	100
754	405	700458687H1	SATMON029	g22485	BLASTN	751	1e-96	95
755	405	700220750H1	SATMON011	g514945	BLASTN	1187	1e-96	96
756	405	700194931H1	SATMON014	g22485	BLASTN	1262	1e-96	99
757	405	700800522H1	SATMON036	g22485	BLASTN	1265	1e-96	98
758	405	700244185H1	SATMON010	g514945	BLASTN	1265	1e-96	100
759	405	700240785H1	SATMON010	g514945	BLASTN	1268	1e-96	98
760	405	700551959H1	SATMON022	g514945	BLASTN	1270	1e-96	100
761	405	700085057H1	SATMON011	g514945	BLASTN	682	1e-95	97
762	405	700332020H1	SATMON019	g514945	BLASTN	713	1e-95	97
763	405	700208841H1	SATMON016	g514945	BLASTN	822	1e-95	95
764	405	700193023H1	SATMON014	g22485	BLASTN	1248	1e-95	98
765	405	700153902H1	SATMON007	g514945	BLASTN	1250	1e-95	100
766	405	700196173H1	SATMON014	g22485	BLASTN	1252	1e-95	99
767	405	700804187H1	SATMON036	g22485	BLASTN	1253	1e-95	98
768	405	700339656H1	SATMON020	g22485	BLASTN	1257	1e-95	99
769	405	700238537H1	SATMON010	g22485	BLASTN	1258	1e-95	99
770	405	700551221H1	SATMON022	g514945	BLASTN	1204	1e-94	99
771	405	700801807H1	SATMON036	g22485	BLASTN	1235	1e-94	100
772	405	700085096H1	SATMON011	g514945	BLASTN	1235	1e-94	100
773	405	700020516H1	SATMON001	g514945	BLASTN	1236	1e-94	98
774	405	700193482H1	SATMON014	g22485	BLASTN	1236	1e-94	99
775	405	700217793H1	SATMON016	g514945	BLASTN	1237	1e-94	98
776	405	700088752H1	SATMON011	g514945	BLASTN	1240	1e-94	100
777	405	700087940H1	SATMON011	g514945	BLASTN	1241	1e-94	99
778	405	700089541H1	SATMON011	g533251	BLASTN	1067	1e-93	91



779	405	700346461H1	SATMON021	g514945	BLASTN	1190	1e-93	99
780	405	700156258H1	SATMON007	g514945	BLASTN	1225	1e-93	100
781	405	700195532H1	SATMON014	g22485	BLASTN	1226	1e-93	99
782	405	700196548H1	SATMON014	g22485	BLASTN	1227	1e-93	99
783	405	700213282H1	SATMON016	g514945	BLASTN	1227	1e-93	96
784	405	700194939H1	SATMON014	g22485	BLASTN	1228	1e-93	98
785	405	700340787H1	SATMON020	g22485	BLASTN	697	1e-92	94
786	405	700105064H1	SATMON010	g22485	BLASTN	706	1e-92	98
787	405	700805132H1	SATMON036	g22485	BLASTN	1046	1e-92	99
788	405	700803414H1	SATMON036	g22485	BLASTN	1211	1e-92	99
789	405	700214594H1	SATMON016	g514945	BLASTN	1212	1e-92	98
790	405	700215080H1	SATMON016	g514945	BLASTN	1213	1e-92	92
791	405	700224219H1	SATMON011	g514945	BLASTN	1213	1e-92	99
792	405	700048190H1	SATMON003	g514945	BLASTN	1218	1e-92	92
793	405	700223284H1	SATMON011	g514945	BLASTN	1220	1e-92	100
794	405	700152413H1	SATMON007	g514945	BLASTN	1220	1e-92	100
795	405	700218003H1	SATMON016	g514945	BLASTN	1221	1e-92	96
796	405	700160155H1	SATMON012	g22485	BLASTN	1222	1e-92	99
797	405	700087943H1	SATMON011	g22485	BLASTN	1222	1e-92	99
798	405	700474049H1	SATMON025	g514945	BLASTN	632	1e-91	97
799	405	700216994H1	SATMON016	g514945	BLASTN	1043	1e-91	99
800	405	700346892H1	SATMON021	g514945	BLASTN	1210	1e-91	96
801	405	700142782H1	SATMON013	g514945	BLASTN	1190	1e-90	100
802	405	700244157H1	SATMON010	g514945	BLASTN	1197	1e-90	97
803	405	700469243H1	SATMON025	g22485	BLASTN	701	1e-89	98
804	405	700618387H1	SATMON033	g514945	BLASTN	853	1e-89	93
805	405	700211154H1	SATMON016	g533251	BLASTN	917	1e-89	90
806	405	700081933H1	SATMON011	g533251	BLASTN	955	1e-89	91
807	405	700235229H1	SATMON010	g514945	BLASTN	955	1e-89	97
808	405	700209241H1	SATMON016	g514945	BLASTN	1076	1e-89	98
809	405	700265841H1	SATMON017	g514945	BLASTN	1097	1e-89	95
810	405	700193804H1	SATMON014	g22485	BLASTN	1180	1e-89	98
811	405	700167742H1	SATMON013	g22485	BLASTN	1180	1e-89	100
812	405	700163256H1	SATMON013	g514945	BLASTN	1182	1e-89	97
813	405	700184973H1	SATMON014	g22485	BLASTN	1182	1e-89	99
814	405	700171785H1	SATMON013	g22485	BLASTN	1184	1e-89	97
815	405	700216915H1	SATMON016	g514945	BLASTN	1185	1e-89	100
816	405	700806685H1	SATMON036	g22485	BLASTN	1186	1e-89	99
817	405	700803846H1	SATMON036	g22485	BLASTN	892	1e-88	95
818	405	700218514H1	SATMON011	g533251	BLASTN	907	1e-88	91
819	405	700574674H1	SATMON030	g22485	BLASTN	957	1e-88	84
820	405	700196082H1	SATMON014	g22485	BLASTN	1054	1e-88	94
821	405	700241637H1	SATMON010	g22485	BLASTN	1081	1e-88	98
822	405	700801876H1	SATMON036	g22485	BLASTN	1163	1e-88	98
823	405	700213443H1	SATMON016	g514945	BLASTN	1171	1e-88	99
824	405	700465181H1	SATMON025	g22485	BLASTN	797	1e-87	92
825	405	700798732H1	SATMON036	g22485	BLASTN	1159	1e-87	96
826	405	700153473H1	SATMON007	g514945	BLASTN	1160	1e-87	100
827	405	700165496H1	SATMON013	g514945	BLASTN	1160	1e-87	97
828	405	700264250H1	SATMON017	g514945	BLASTN	643	1e-86	100
829	405	700335490H1	SATMON019	g514945	BLASTN	671	1e-86	97
830	405	700575891H1	SATMON030	g22485	BLASTN	780	1e-86	92
831	405	700222931H1	SATMON011	g514945	BLASTN	1117	1e-86	91
832	405	700163588H1	SATMON013	g514945	BLASTN	1140	1e-86	100

833	405	700161111H1	SATMON012	g22485	BLASTN	1142	1e-86	99
834	405	700016023H1	SATMON001	g514945	BLASTN	1147	1e-86	99
835	405	700209043H1	SATMON016	g514945	BLASTN	659	1e-85	98
836	405	700333556H1	SATMON019	g514945	BLASTN	783	1e-85	89
837	405	700570471H1	SATMON030	g22485	BLASTN	831	1e-85	89
838	405	700171040H1	SATMON013	g22485	BLASTN	1127	1e-85	99
839	405	700196146H1	SATMON014	g22485	BLASTN	1128	1e-85	96
840	405	700021837H1	SATMON001	g514945	BLASTN	1130	1e-85	98
841	405	700169062H1	SATMON013	g514945	BLASTN	1130	1e-85	100
842	405	700218263H1	SATMON016	g514945	BLASTN	1135	1e-85	100
843	405	700091281H1	SATMON011	g514945	BLASTN	1135	1e-85	100
844	405	700221221H1	SATMON011	g22485	BLASTN	735	1e-84	98
845	405	700239933H1	SATMON010	g22485	BLASTN	779	1e-84	97
846	405	700193089H1	SATMON014	g22485	BLASTN	979	1e-84	97
847	405	700216767H1	SATMON016	g514945	BLASTN	1005	1e-84	98
848	405	700085786H1	SATMON011	g514945	BLASTN	1117	1e-84	99
849	405	700163868H1	SATMON013	g22485	BLASTN	1120	1e-84	100
850	405	700167028H1	SATMON013	g514945	BLASTN	1120	1e-84	100
851	405	700170416H1	SATMON013	g514945	BLASTN	1120	1e-84	100
852	405	700266595H1	SATMON017	g514945	BLASTN	1120	1e-84	100
853	405	700377630H1	SATMON019	g514945	BLASTN	649	1e-83	95
854	405	700207830H1	SATMON016	g514945	BLASTN	877	1e-83	97
855	405	700241726H1	SATMON010	g514945	BLASTN	1104	1e-83	97
856	405	700806447H1	SATMON036	g22485	BLASTN	1106	1e-83	93
857	405	700018166H1	SATMON001	g514945	BLASTN	1108	1e-83	98
858	405	700083463H1	SATMON011	g514945	BLASTN	633	1e-82	92
859	405	700548890H1	SATMON022	g22485	BLASTN	727	1e-82	93
860	405	700218553H1	SATMON011	g22485	BLASTN	979	1e-82	95
861	405	700016408H1	SATMON001	g514945	BLASTN	1026	1e-82	97
862	405	700569427H2	SATMON030	g514945	BLASTN	1095	1e-82	97
863	405	700172546H1	SATMON013	g514945	BLASTN	1100	1e-82	100
864	405	700804387H1	SATMON036	g22485	BLASTN	668	1e-81	97
865	405	700155788H1	SATMON007	g514945	BLASTN	840	1e-81	100
866	405	700807167H1	SATMON036	g22485	BLASTN	1024	1e-81	97
867	405	700472356H1	SATMON025	g22485	BLASTN	1080	1e-81	98
868	405	700193535H1	SATMON014	g22485	BLASTN	1082	1e-81	99
869	405	700171177H1	SATMON013	g22485	BLASTN	1086	1e-81	98
870	405	700799867H1	SATMON036	g22485	BLASTN	1087	1e-81	96
871	405	700263716H1	SATMON017	g514945	BLASTN	1089	1e-81	92
872	405	700476045H1	SATMON025	g22485	BLASTN	608	1e-80	88
873	405	700803344H1	SATMON036	g22485	BLASTN	834	1e-80	97
874	405	700168924H1	SATMON013	g514945	BLASTN	860	1e-80	99
875	405	700218569H1	SATMON011	g22485	BLASTN	900	1e-80	98
876	405	700088574H1	SATMON011	g514945	BLASTN	900	1e-80	86
877	405	700471932H1	SATMON025	g530978	BLASTN	1064	1e-80	83
878	405	700020011H1	SATMON001	g22485	BLASTN	1067	1e-80	99
879	405	700167511H1	SATMON013	g22485	BLASTN	1070	1e-80	100
880	405	700219249H1	SATMON011	g514945	BLASTN	1070	1e-80	100
881	405	700804846H1	SATMON036	g22485	BLASTN	1075	1e-80	90
882	405	700150388H1	SATMON007	g22485	BLASTN	1075	1e-80	100
883	405	700807395H1	SATMON036	g22485	BLASTN	571	1e-79	90
884	405	700090864H1	SATMON011	g514945	BLASTN	630	1e-79	100
885	405	700217812H1	SATMON016	g514945	BLASTN	646	1e-79	91
886	405	700203618H1	SATMON003	g22485	BLASTN	913	1e-79	96

887	405	700203302H1	SATMON003	g514945	BLASTN	1030	1e-79	100
888	405	700163192H1	SATMON013	g22485	BLASTN	1056	1e-79	97
889	405	700805065H1	SATMON036	g22485	BLASTN	1066	1e-79	95
890	405	700086763H1	SATMON011	g514945	BLASTN	901	1e-78	98
891	405	700240070H1	SATMON010	g533251	BLASTN	923	1e-78	90
892	405	700018847H1	SATMON001	g22485	BLASTN	1045	1e-78	98
893	405	700803420H1	SATMON036	g22485	BLASTN	1048	1e-78	96
894	405	700799936H1	SATMON036	g22485	BLASTN	1050	1e-78	96
895	405	700207637H1	SATMON016	g514945	BLASTN	828	1e-77	98
896	405	700807034H1	SATMON036	g22485	BLASTN	808	1e-76	91
897	405	700198035H1	SATMON016	g514945	BLASTN	1025	1e-76	100
898	405	700169076H1	SATMON013	g514945	BLASTN	1028	1e-76	99
899	405	700020564H1	SATMON001	g514945	BLASTN	1030	1e-76	98
900	405	700799968H1	SATMON036	g22485	BLASTN	701	1e-75	99
901	405	700378056H1	SATMON019	g22485	BLASTN	802	1e-75	97
902	405	700219691H1	SATMON011	g514945	BLASTN	1018	1e-75	99
903	405	700168945H1	SATMON013	g22485	BLASTN	848	1e-74	94
904	405	700242730H1	SATMON010	g514945	BLASTN	1006	1e-74	99
905	405	700210096H1	SATMON016	g514945	BLASTN	756	1e-73	93
906	405	700333941H1	SATMON019	g514945	BLASTN	923	1e-73	99
907	405	700576645H1	SATMON030	g22485	BLASTN	991	1e-73	99
908	405	700333494H1	SATMON019	g514945	BLASTN	601	1e-72	91
909	405	700023296H1	SATMON003	g514945	BLASTN	726	1e-72	95
910	405	700802508H1	SATMON036	g22485	BLASTN	811	1e-72	94
911	405	700223382H1	SATMON011	g22485	BLASTN	865	1e-72	98
912	405	700215535H1	SATMON016	g514945	BLASTN	942	1e-72	96
913	405	700017549H1	SATMON001	g514945	BLASTN	973	1e-72	97
914	405	700799113H1	SATMON036	g22485	BLASTN	787	1e-70	99
915	405	700168696H1	SATMON013	g514945	BLASTN	946	1e-69	89
916	405	700088269H1	SATMON011	g514945	BLASTN	946	1e-69	93
917	405	700194522H1	SATMON014	g22485	BLASTN	875	1e-68	97
918	405	700203476H1	SATMON003	g22485	BLASTN	923	1e-68	86
919	405	700549205H1	SATMON022	g22485	BLASTN	300	1e-66	89
920	405	700196217H1	SATMON014	g22485	BLASTN	907	1e-66	96
921	405	700163647H1	SATMON013	g22485	BLASTN	888	1e-65	98
922	405	700804485H1	SATMON036	g22485	BLASTN	896	1e-65	99
923	405	700193074H1	SATMON014	g22485	BLASTN	605	1e-63	96
924	405	700203370H1	SATMON003	g514945	BLASTN	857	1e-62	98
925	405	700201575H1	SATMON003	g514945	BLASTN	335	1e-60	87
926	405	700378020H1	SATMON019	g514945	BLASTN	833	1e-60	97
927	405	700242865H1	SATMON010	g514945	BLASTN	823	1e-59	91
928	405	700344036H1	SATMON021	g514945	BLASTN	825	1e-59	100
929	405	700215849H1	SATMON016	g514945	BLASTN	805	1e-58	100
930	405	700443538H1	SATMON027	g22485	BLASTN	814	1e-58	98
931	405	700804448H1	SATMON036	g22485	BLASTN	791	1e-57	99
932	405	700155008H1	SATMON007	g22485	BLASTN	802	1e-57	98
933	405	700201244H1	SATMON003	g22485	BLASTN	530	1e-56	97
934	405	700616378H1	SATMON033	g22485	BLASTN	682	1e-56	97
935	405	700333357H1	SATMON019	g22485	BLASTN	780	1e-56	80
936	405	700222360H1	SATMON011	g514945	BLASTN	777	1e-55	92
937	405	700214724H1	SATMON016	g514945	BLASTN	763	1e-54	98
938	405	700571283H1	SATMON030	g514945	BLASTN	736	1e-52	99
939	405	700020194H1	SATMON001	g22485	BLASTN	415	1e-51	99
940	405	700620551H1	SATMON034	g22485	BLASTN	473	1e-51	95

941	405	700446320H1	SATMON027	g22485	BLASTN	475	1e-50	87
942	405	700241357H1	SATMON010	g22485	BLASTN	701	1e-49	99
943	405	700617094H1	SATMON033	g22485	BLASTN	673	1e-47	97
944	405	700206691H1	SATMON003	g514945	BLASTN	680	1e-47	90
945	405	700091580H1	SATMON011	g514945	BLASTN	680	1e-47	100
946	405	700574515H1	SATMON030	g514945	BLASTN	369	1e-46	74
947	405	700155148H1	SATMON007	g514945	BLASTN	397	1e-45	97
948	405	700612388H1	SATMON033	g514945	BLASTN	625	1e-43	100
949	405	700474681H1	SATMON025	g22485	BLASTN	379	1e-41	91
950	405	700800401H1	SATMON036	g22485	BLASTN	395	1e-40	90
951	405	700155657H1	SATMON007	g514945	BLASTN	591	1e-40	95
952	405	700076002H1	SATMON007	g514945	BLASTN	575	1e-39	100
953	405	700802090H1	SATMON036	g22485	BLASTN	577	1e-39	98
954	405	700170104H1	SATMON013	g22485	BLASTN	565	1e-38	100
955	405	701183763H1	SATMONN06	g514945	BLASTN	569	1e-38	90
956	405	700084688H1	SATMON011	g514945	BLASTN	380	1e-36	98
957	405	700473655H1	SATMON025	g22485	BLASTN	530	1e-35	100
958	405	700615166H1	SATMON033	g514945	BLASTN	531	1e-35	94
959	405	700085562H1	SATMON011	g533251	BLASTN	532	1e-35	98
960	405	700153049H1	SATMON007	g514945	BLASTN	537	1e-35	94
961	405	700090656H1	SATMON011	g514945	BLASTN	489	1e-34	98
962	405	700802054H1	SATMON036	g22485	BLASTN	345	1e-31	99
963	405	700802284H1	SATMON036	g22485	BLASTN	488	1e-31	97
964	405	700802312H1	SATMON036	g22485	BLASTN	270	1e-30	100
965	405	700153683H1	SATMON007	g22485	BLASTN	461	1e-29	98
966	405	700028453H1	SATMON003	g22485	BLASTN	321	1e-27	99
967	405	700089391H1	SATMON011	g514945	BLASTN	404	1e-24	96
968	405	700381969H1	SATMON023	g22485	BLASTN	385	1e-23	94
969	405	700800135H1	SATMON036	g22485	BLASTN	180	1e-21	100
970	405	700088173H1	SATMON011	g514945	BLASTN	347	1e-20	95
971	405	700202170H1	SATMON003	g19108	BLASTX	133	1e-11	96
972	537	700209929H1	SATMON016	g22485	BLASTN	1478	1e-114	99
973	537	700096948H1	SATMON008	g22485	BLASTN	911	1e-113	99
974	537	700476287H1	SATMON025	g22485	BLASTN	1403	1e-108	98
975	537	700803088H1	SATMON036	g22485	BLASTN	1336	1e-107	96
976	537	700799436H1	SATMON036	g22485	BLASTN	1361	1e-104	99
977	537	700224822H1	SATMON011	g22485	BLASTN	1300	1e-103	96
978	537	700241134H1	SATMON010	g22485	BLASTN	1302	1e-99	99
979	537	700803625H1	SATMON036	g22485	BLASTN	1292	1e-98	99
980	537	700802549H1	SATMON036	g22485	BLASTN	1232	1e-93	99
981	537	700477992H1	SATMON025	g22485	BLASTN	943	1e-92	97
982	537	700150953H1	SATMON007	g22485	BLASTN	1152	1e-87	99
983	537	700205638H1	SATMON003	g22485	BLASTN	1086	1e-81	99
984	537	700803732H1	SATMON036	g22487	BLASTN	379	1e-79	97
985	537	700165461H1	SATMON013	g22485	BLASTN	1064	1e-79	98
986	537	700807069H1	SATMON036	g22485	BLASTN	957	1e-77	96
987	537	700800902H1	SATMON036	g22485	BLASTN	762	1e-54	86
988	537	700466671H1	SATMON025	g22485	BLASTN	520	1e-44	95
989	537	700799118H1	SATMON036	g22485	BLASTN	626	1e-43	99
990	537	700802273H1	SATMON036	g22485	BLASTN	616	1e-42	99
991	537	700804848H1	SATMON036	g22485	BLASTN	306	1e-33	98
992	8549	700103190H1	SATMON010	g1100216	BLASTN	615	1e-92	98
993	8549	700075574H1	SATMON007	g1100216	BLASTN	701	1e-92	100
994	8549	700218547H1	SATMON011	g514945	BLASTN	1208	1e-91	99

995	8549	700213873H1	SATMON016	g1100216	BLASTN	673	1e-90	95
996	8549	700221147H1	SATMON011	g1100216	BLASTN	646	1e-89	98
997	8549	700207093H1	SATMON003	g1100216	BLASTN	701	1e-87	100
998	8549	700210112H1	SATMON016	g1100216	BLASTN	615	1e-84	98
999	8549	700096984H1	SATMON008	g514945	BLASTN	1111	1e-83	99
1000	8549	700221070H1	SATMON011	g1100216	BLASTN	645	1e-82	96
1001	8549	700332046H1	SATMON019	g1100216	BLASTN	601	1e-76	89
1002	8549	700150377H1	SATMON007	g1100216	BLASTN	621	1e-74	100
1003	8549	700084780H1	SATMON011	g514945	BLASTN	585	1e-39	100
1004	8549	700153082H1	SATMON007	g1100216	BLASTN	495	1e-36	89
1005	8549	700261144H1	SATMON017	g1100216	BLASTN	339	1e-35	87
1006	8549	700264112H1	SATMON017	g1100216	BLASTN	428	1e-34	91
1007	8549	700473660H1	SATMON025	g1100216	BLASTN	415	1e-28	100
1008	8549	700473628H1	SATMON025	g514945	BLASTN	329	1e-26	88
1009	8549	700351060H1	SATMON023	g1100216	BLASTN	291	1e-22	91
1010	-L30595280	LIB3059-039-Q1-K1-A5	LIB3059	g22485	BLASTN	473	1e-30	79
1011	-L30612133	LIB3061-024-Q1-K1-H5	LIB3061	g22485	BLASTN	849	1e-61	80
1012	-L30616296	LIB3061-043-Q1-K1-A10	LIB3061	g22485	BLASTN	479	1e-98	82
1013	-L30623037	LIB3062-030-Q1-K1-F12	LIB3062	g514945	BLASTN	684	1e-48	78
1014	-L30625289	LIB3062-021-Q1-K1-C2	LIB3062	g514945	BLASTN	1180	1e-111	79
1015	-L30663565	LIB3066-053-Q1-K1-D6	LIB3066	g530978	BLASTN	568	1e-36	76
1016	-L30784420	LIB3078-039-Q1-K1-A4	LIB3078	g514945	BLASTN	484	1e-40	81
1017	30444	LIB3069-052-Q1-K1-F8	LIB3069	g1100216	BLASTN	558	1e-77	89
1018	32909	LIB143-057-Q1-E1-F6	LIB143	g2570066	BLASTN	902	1e-69	74
1019	405	LIB3062-021-Q1-K1-C5	LIB3062	g514945	BLASTN	2368	1e-188	99
1020	405	LIB3078-024-Q1-K1-C5	LIB3078	g514945	BLASTN	2356	1e-187	98
1021	405	LIB3059-028-Q1-K1-D5	LIB3059	g22485	BLASTN	2163	1e-171	98
1022	405	LIB3059-015-Q1-K1-E7	LIB3059	g22485	BLASTN	2167	1e-171	98
1023	405	LIB3059-044-Q1-K1-E7	LIB3059	g514945	BLASTN	2170	1e-171	98
1024	405	LIB3061-029-Q1-K1-G11	LIB3061	g22485	BLASTN	2055	1e-170	98
1025	405	LIB3059-011-Q1-K1-F5	LIB3059	g22485	BLASTN	2137	1e-169	98
1026	405	LIB3062-009-Q1-K1-D1	LIB3062	g514945	BLASTN	2122	1e-167	98
1027	405	LIB3061-011-Q1-K1-D9	LIB3061	g22485	BLASTN	2091	1e-165	98
1028	405	LIB3067-040-Q1-K1-E8	LIB3067	g514945	BLASTN	1916	1e-164	99
1029	405	LIB3062-041-	LIB3062	g514945	BLASTN	2082	1e-164	97

1030	405	Q1-K1-D4 LIB3062-022-	LIB3062	g514945	BLASTN	2084	1e-164	99
1031	405	Q1-K1-C9 LIB3062-033-	LIB3062	g514945	BLASTN	1854	1e-161	95
1032	405	Q1-K1-C7 LIB3062-002-	LIB3062	g514945	BLASTN	1854	1e-161	97
1033	405	Q1-K2-F9 LIB3059-010-	LIB3059	g22485	BLASTN	2018	1e-159	99
1034	405	Q1-K1-C9 LIB3059-013-	LIB3059	g22485	BLASTN	2022	1e-159	98
1035	405	Q1-K1-B10 LIB3061-020-	LIB3061	g22485	BLASTN	1771	1e-158	97
1036	405	Q1-K1-F2 LIB3061-022-	LIB3061	g22485	BLASTN	1909	1e-158	98
1037	405	Q1-K1-C2 LIB3062-023-	LIB3062	g22485	BLASTN	1508	1e-157	96
1038	405	Q1-K1-D10 LIB3061-008-	LIB3061	g22485	BLASTN	1983	1e-156	97
1039	405	Q1-K1-H11 LIB3059-024-	LIB3059	g22485	BLASTN	1051	1e-154	99
1040	405	Q1-K1-H4 LIB3062-048-	LIB3062	g22485	BLASTN	1187	1e-154	94
1041	405	Q1-K1-G5 LIB3061-025-	LIB3061	g22485	BLASTN	1803	1e-154	95
1042	405	Q1-K1-B1 LIB3061-028-	LIB3061	g22485	BLASTN	1963	1e-154	97
1043	405	Q1-K1-C4 LIB3078-057-	LIB3078	g514945	BLASTN	1412	1e-153	92
1044	405	Q1-K1-D9 LIB3061-021-	LIB3061	g22485	BLASTN	1465	1e-153	96
1045	405	Q1-K1-A8 LIB3061-025-	LIB3061	g22485	BLASTN	1524	1e-153	96
1046	405	Q1-K1-B5 LIB3061-008-	LIB3061	g22485	BLASTN	1879	1e-153	94
1047	405	Q1-K1-C7 LIB3061-039-	LIB3061	g22485	BLASTN	1853	1e-151	96
1048	405	Q1-K1-A8 LIB3078-039-	LIB3078	g514945	BLASTN	1853	1e-151	96
1049	405	Q1-K1-E5 LIB3061-049-	LIB3061	g22485	BLASTN	1801	1e-150	98
1050	405	Q1-K2-G2 LIB3062-001-	LIB3062	g514945	BLASTN	1916	1e-150	94
1051	405	Q1-K1-G6 LIB3061-021-	LIB3061	g22485	BLASTN	1918	1e-150	92
1052	405	Q1-K1-D2 LIB3061-039-	LIB3061	g22485	BLASTN	1361	1e-149	96
1053	405	Q1-K1-G8 LIB3061-051-	LIB3061	g22485	BLASTN	1768	1e-148	98
1054	405	Q1-K1-A12 LIB3061-015-	LIB3061	g22485	BLASTN	1667	1e-146	93
1055	405	Q1-K1-H11 LIB3059-040-	LIB3059	g22485	BLASTN	1835	1e-146	97
1056	405	Q1-K2-G5 LIB3061-002-	LIB3061	g22485	BLASTN	1845	1e-144	89
1057	405	LIB3062-002-	LIB3062	g22485	BLASTN	1672	1e-142	99

		Q1-K2-G12						
1057	405	LIB3059-048-Q1-K1-H5	LIB3059	g22485	BLASTN	1822	1e-142	99
1058	405	LIB3078-040-Q1-K1-F8	LIB3078	g514945	BLASTN	1801	1e-141	97
1059	405	LIB3078-001-Q1-K1-C7	LIB3078	g22485	BLASTN	1246	1e-139	95
1060	405	LIB3061-024-Q1-K1-A12	LIB3061	g22485	BLASTN	1376	1e-139	94
1061	405	LIB3061-026-Q1-K1-D3	LIB3061	g22485	BLASTN	1643	1e-138	93
1062	405	LIB3061-056-Q1-K1-D8	LIB3061	g22485	BLASTN	1763	1e-138	92
1063	405	LIB3069-041-Q1-K1-G12	LIB3069	g514945	BLASTN	1758	1e-137	97
1064	405	LIB3059-025-Q1-K1-E5	LIB3059	g22485	BLASTN	1532	1e-132	94
1065	405	LIB3061-014-Q1-K1-D4	LIB3061	g22485	BLASTN	1294	1e-130	88
1066	405	LIB3061-005-Q1-K1-C9	LIB3061	g22485	BLASTN	1540	1e-130	97
1067	405	LIB3061-016-Q1-K1-G2	LIB3061	g22485	BLASTN	1251	1e-129	85
1068	405	LIB3069-029-Q1-K1-B2	LIB3069	g514945	BLASTN	1657	1e-129	88
1069	405	LIB3078-012-Q1-K1-F7	LIB3078	g514945	BLASTN	857	1e-128	86
1070	405	LIB3078-016-Q1-K1-D7	LIB3078	g514945	BLASTN	1335	1e-128	87
1071	405	LIB3062-049-Q1-K1-A8	LIB3062	g514945	BLASTN	1609	1e-128	88
1072	405	LIB143-006-Q1-E1-G12	LIB143	g514945	BLASTN	1614	1e-125	96
1073	405	LIB3059-024-Q1-K1-E5	LIB3059	g22485	BLASTN	1529	1e-123	83
1074	405	LIB3069-008-Q1-K1-C1	LIB3069	g514945	BLASTN	1036	1e-115	94
1075	405	LIB3059-018-Q1-K1-F11	LIB3059	g514945	BLASTN	910	1e-103	93
1076	405	LIB3078-001-Q1-K1-E8	LIB3078	g514945	BLASTN	952	1e-98	90
1077	405	LIB3059-017-Q1-K1-G4	LIB3059	g22485	BLASTN	1170	1e-88	92
1078	405	LIB3067-045-Q1-K1-E9	LIB3067	g533251	BLASTN	917	1e-87	87
1079	405	LIB3062-015-Q1-K1-C1	LIB3062	g514945	BLASTN	1066	1e-86	96
1080	405	LIB3059-039-Q1-K1-A3	LIB3059	g22485	BLASTN	856	1e-82	92
1081	405	LIB3062-024-Q1-K1-C3	LIB3062	g514945	BLASTN	548	1e-79	88
1082	405	LIB3059-029-Q1-K1-F1	LIB3059	g22485	BLASTN	925	1e-74	94
1083	405	LIB3059-006-	LIB3059	g22485	BLASTN	530	1e-50	83

1084	405	Q1-K1-F4 LIB3067-017-	LIB3067	g533251	BLASTN	425	1e-26	100
1085	405	Q1-K1-C3 LIB3061-028-	LIB3061	g19106	BLASTX	118	1e-25	100
1086	537	Q1-K1-A9 LIB3066-009-	LIB3066	g22485	BLASTN	1369	1e-122	96
		Q1-K1-B9						

MAIZE HEXOKINASE								
Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
1087	-700018381	700018381H1	SATMON001	g1899025	BLASTX	166	1e-16	48
1088	-700051079	700051079H1	SATMON003	g1899025	BLASTX	84	1e-11	50
1089	-700101579	700101579H1	SATMON009	g881521	BLASTX	217	1e-23	66
1090	-700105594	700105594H1	SATMON010	g3087888	BLASTX	181	1e-17	57
1091	-700106018	700106018H1	SATMON010	g3087888	BLASTX	195	1e-19	64
1092	-700157233	700157233H1	SATMON012	g3087888	BLASTX	198	1e-20	58
1093	-700202992	700202992H1	SATMON003	g3087888	BLASTX	89	1e-9	58
1094	-700224204	700224204H1	SATMON011	g1899024	BLASTN	520	1e-34	70
1095	-700241273	700241273H1	SATMON010	g3087888	BLASTX	184	1e-18	58
1096	-700352183	700352183H1	SATMON023	g1899024	BLASTN	481	1e-31	70
1097	-700573814	700573814H1	SATMON030	g1899024	BLASTN	535	1e-34	67
1098	-700612458	700612458H1	SATMON033	g619928	BLASTX	229	1e-26	61
1099	-701168774	701168774H1	SATMONN05	g619927	BLASTN	252	1e-10	62
1100	1195	700457430H1	SATMON029	g3087888	BLASTX	122	1e-19	53
1101	13262	700102942H1	SATMON010	g3087888	BLASTX	113	1e-18	53
1102	1378	700456148H1	SATMON029	g1899025	BLASTX	267	1e-29	59
1103	1378	700455837H1	SATMON029	g1899025	BLASTX	166	1e-21	60
1104	17305	700460742H1	SATMON031	g619928	BLASTX	131	1e-15	57
1105	17305	700614972H1	SATMON033	g1899025	BLASTX	100	1e-8	53
1106	1842	700089135H1	SATMON011	g619928	BLASTX	405	1e-49	70
1107	1842	700430234H1	SATMONN01	g619927	BLASTN	461	1e-28	72
1108	1842	700166122H1	SATMON013	g619928	BLASTX	183	1e-18	84
1109	24376	700053677H1	SATMON010	g1899024	BLASTN	642	1e-44	70
1110	24376	700152328H1	SATMON007	g619927	BLASTN	555	1e-37	69
1111	24376	700623451H1	SATMON034	g619928	BLASTX	197	1e-32	72
1112	28388	700089065H1	SATMON011	g619928	BLASTX	186	1e-30	61
1113	3345	700072110H1	SATMON007	g619928	BLASTX	125	1e-24	66
1114	3345	700472061H1	SATMON025	g619928	BLASTX	112	1e-20	55
1115	3345	701173753H1	SATMONN05	g619928	BLASTX	135	1e-16	54
1116	3345	700202130H1	SATMON003	g619928	BLASTX	113	1e-11	68
1117	5073	700582054H1	SATMON031	g619928	BLASTX	247	1e-29	66
1118	5073	700053432H1	SATMON009	g619928	BLASTX	233	1e-25	60
1119	6731	700099009H1	SATMON009	g619927	BLASTN	736	1e-52	72
1120	6731	700089738H1	SATMON011	g1899024	BLASTN	700	1e-49	70
1121	6731	700171542H1	SATMON013	g619927	BLASTN	530	1e-35	74
1122	7565	700356773H1	SATMON024	g1899025	BLASTX	177	1e-17	62
1123	9695	700212172H1	SATMON016	g1899024	BLASTN	832	1e-60	74
1124	9695	700212124H1	SATMON016	g1899024	BLASTN	835	1e-60	75
1125	9695	700094278H1	SATMON008	g1899024	BLASTN	819	1e-59	74
1126	-L30621307	LIB3062-001-	LIB3062	g1899025	BLASTX	95	1e-32	53
		Q1-K2-G11						
1127	-L30782665	LIB3078-007-	LIB3078	g3087888	BLASTX	130	1e-39	47
		Q1-K1-E9						



1128	24376	LIB3069-041-Q1-K1-E7	LIB3069	g1899024	BLASTN	608	1e-61	70
1129	28244	LIB3061-004-Q1-K1-F9	LIB3061	g687676	BLASTN	499	1e-30	65
1130	28388	LIB3066-030-Q1-K1-G10	LIB3066	g619928	BLASTX	299	1e-63	64
1131	3364	LIB3078-051-Q1-K1-B3	LIB3078	g687676	BLASTN	619	1e-41	67
1132	3364	LIB3078-053-Q1-K1-C9	LIB3078	g687676	BLASTN	627	1e-41	69
1133	3364	LIB84-015-Q1-E1-F7	LIB84	g687676	BLASTN	554	1e-35	69
1134	6731	LIB3061-028-Q1-K1-C1	LIB3061	g1899024	BLASTN	831	1e-60	70
1135	9695	LIB143-065-Q1-E1-C10	LIB143	g1899024	BLASTN	1096	1e-82	73

#### MAIZE FRUCTOKINASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
1136	-700106058	700106058H1	SATMON010	g1052972	BLASTN	220	1e-9	68
1137	-700151135	700151135H1	SATMON007	g297014	BLASTN	351	1e-18	75
1138	-700169310	700169310H1	SATMON013	g1052972	BLASTN	273	1e-12	59
1139	-700210226	700210226H1	SATMON016	g1052973	BLASTX	188	1e-24	68
1140	-700257901	700257901H1	SATMON017	g297015	BLASTX	200	1e-20	72
1141	-700621274	700621274H1	SATMON034	g1052973	BLASTX	141	1e-24	64
1142	11678	700105513H1	SATMON010	g1052972	BLASTN	580	1e-39	64
1143	11678	700170725H1	SATMON013	g1052972	BLASTN	478	1e-31	66
1144	2526	700159958H1	SATMON012	g1052973	BLASTX	152	1e-14	64
1145	2754	700102678H1	SATMON010	g1052972	BLASTN	707	1e-50	69
1146	2754	700102312H1	SATMON010	g1052972	BLASTN	701	1e-49	69
1147	2754	700205695H1	SATMON003	g1915973	BLASTN	633	1e-43	69
1148	2754	700221511H1	SATMON011	g1915973	BLASTN	587	1e-40	69
1149	2754	700469079H1	SATMON025	g1052972	BLASTN	584	1e-39	72
1150	2754	701173520H1	SATMONN05	g1915973	BLASTN	342	1e-36	70
1151	2754	700267332H1	SATMON017	g1052972	BLASTN	541	1e-35	64
1152	2754	701164907H1	SATMONN04	g1052973	BLASTX	280	1e-33	57
1153	2754	700450050H2	SATMON028	g1052973	BLASTX	160	1e-31	60
1154	2754	701182860H1	SATMONN06	g297015	BLASTX	188	1e-27	65
1155	2754	700467520H1	SATMON025	g1915974	BLASTX	242	1e-26	60
1156	2754	700159848H1	SATMON012	g1052973	BLASTX	197	1e-24	63
1157	3287	700088103H1	SATMON011	g2102693	BLASTX	239	1e-43	74
1158	3287	700210913H1	SATMON016	g2102693	BLASTX	250	1e-35	77
1159	3287	700167609H1	SATMON013	g1052973	BLASTX	300	1e-35	68
1160	3287	700085916H1	SATMON011	g1052972	BLASTN	553	1e-35	64
1161	3287	700262715H1	SATMON017	g1915974	BLASTX	201	1e-33	71
1162	3287	700170179H1	SATMON013	g1052973	BLASTX	289	1e-33	67
1163	3287	700615671H1	SATMON033	g1052972	BLASTN	515	1e-32	63
1164	3287	700223640H1	SATMON011	g1052973	BLASTX	219	1e-31	67
1165	3287	700215234H1	SATMON016	g1052973	BLASTX	190	1e-30	67
1166	3287	700203946H1	SATMON003	g1052973	BLASTX	198	1e-30	60
1167	3287	700028411H1	SATMON003	g2102693	BLASTX	110	1e-29	57
1168	3287	700224307H1	SATMON011	g1052973	BLASTX	159	1e-29	87
1169	3287	700072013H1	SATMON007	g1052973	BLASTX	191	1e-29	65

1170	3287	700215669H1	SATMON016	g1052973	BLASTX	260	1e-29	57
1171	3287	700353954H1	SATMON024	g1052973	BLASTX	260	1e-29	61
1172	3287	700342211H1	SATMON021	g1052973	BLASTX	137	1e-28	67
1173	3287	700085462H1	SATMON011	g297014	BLASTN	466	1e-28	62
1174	3287	700220972H1	SATMON011	g1052973	BLASTX	109	1e-27	83
1175	3287	700451141H1	SATMON028	g1052973	BLASTX	245	1e-27	63
1176	3287	700087484H1	SATMON011	g1052972	BLASTN	440	1e-26	64
1177	3287	700343411H1	SATMON021	g1052973	BLASTX	163	1e-25	67
1178	3287	700217263H1	SATMON016	g1915973	BLASTN	393	1e-25	68
1179	3287	700030665H1	SATMON003	g1052973	BLASTX	176	1e-24	71
1180	3287	700343380H1	SATMON021	g1052973	BLASTX	228	1e-24	57
1181	3287	701159743H2	SATMONN04	g1052973	BLASTX	183	1e-23	55
1182	3287	700221543H1	SATMON011	g1052973	BLASTX	217	1e-23	50
1183	3287	700333946H1	SATMON019	g1052973	BLASTX	178	1e-22	66
1184	3287	700091730H1	SATMON011	g1052973	BLASTX	171	1e-21	64
1185	3287	700570521H1	SATMON030	g1915974	BLASTX	98	1e-18	58
1186	3287	700048604H1	SATMON003	g1052973	BLASTX	88	1e-15	54
1187	3287	700208681H1	SATMON016	g1052973	BLASTX	129	1e-15	55
1188	3287	700028328H1	SATMON003	g1052973	BLASTX	162	1e-15	66
1189	3287	700220530H1	SATMON011	g1052973	BLASTX	141	1e-14	88
1190	3287	700243726H1	SATMON010	g1052973	BLASTX	153	1e-14	68
1191	3287	700142502H1	SATMON012	g1052973	BLASTX	157	1e-14	47
1192	3287	700336537H1	SATMON019	g1052973	BLASTX	141	1e-12	50
1193	3287	700205308H1	SATMON003	g1052973	BLASTX	133	1e-11	75
1194	5966	700084171H1	SATMON011	g1052972	BLASTN	448	1e-26	66
1195	5966	700084951H1	SATMON011	g2102693	BLASTX	214	1e-22	73
1196	5966	700089353H1	SATMON011	g2102691	BLASTX	195	1e-20	72
1197	5966	700220723H1	SATMON011	g1915974	BLASTX	198	1e-20	73
1198	5966	700084412H1	SATMON011	g2102693	BLASTX	179	1e-19	76
1199	5966	700085628H1	SATMON011	g2102691	BLASTX	180	1e-18	72
1200	5966	700027982H1	SATMON003	g2102691	BLASTX	178	1e-17	72
1201	5966	700106884H1	SATMON010	g1915974	BLASTX	148	1e-13	75
1202	5966	700053135H1	SATMON008	g1915974	BLASTX	131	1e-11	73
1203	5966	700027988H1	SATMON003	g1915974	BLASTX	134	1e-11	65
1204	5966	700207083H1	SATMON003	g1915974	BLASTX	100	1e-10	46
1205	5966	700158574H1	SATMON012	g1915974	BLASTX	120	1e-9	50
1206	2754	LIB3061-030-Q1-K1-G12	LIB3061	g1052972	BLASTN	882	1e-64	67
1207	2754	LIB3061-030-Q1-K1-G11	LIB3061	g1052972	BLASTN	751	1e-52	68
1208	3287	LIB3067-040-Q1-K1-H10	LIB3067	g1052972	BLASTN	657	1e-44	64
1209	3287	LIB84-024-Q1-E1-H7	LIB84	g1052972	BLASTN	638	1e-42	64
1210	3287	LIB3069-045-Q1-K1-F6	LIB3069	g1052972	BLASTN	592	1e-38	61
1211	3287	LIB3061-014-Q1-K1-A3	LIB3061	g1052973	BLASTX	175	1e-36	41
1212	3287	LIB3062-019-Q1-K1-H11	LIB3062	g1052973	BLASTX	154	1e-30	68
1213	3287	LIB3067-054-Q1-K1-C9	LIB3067	g1052972	BLASTN	495	1e-30	61
1214	3287	LIB3067-022-Q1-K1-H4	LIB3067	g1052973	BLASTX	141	1e-27	68

1215	3287	LIB3069-045-Q1-K1-F2	LIB3069	g1052972	BLASTN	439	1e-25	57
------	------	----------------------	---------	----------	--------	-----	-------	----

# MAIZE NDP-KINASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
1216	-700575072	700575072H1	SATMON030	g303849	BLASTX	74	1e-13	89
1217	-701170773	701170773H1	SATMONN05	g1777930	BLASTX	132	1e-30	71
1218	2462	700050003H1	SATMON003	g218233	BLASTN	656	1e-58	83
1219	2462	700204789H1	SATMON003	g218233	BLASTN	780	1e-58	87
1220	2462	700049819H1	SATMON003	g218233	BLASTN	786	1e-58	86
1221	2462	700204211H1	SATMON003	g218233	BLASTN	786	1e-58	86
1222	2462	700205742H1	SATMON003	g218233	BLASTN	763	1e-57	86
1223	2462	700207611H1	SATMON016	g218233	BLASTN	764	1e-57	87
1224	2462	700072505H1	SATMON007	g218233	BLASTN	740	1e-55	86
1225	2462	700236468H1	SATMON010	g218233	BLASTN	710	1e-52	86
1226	2462	701181270H1	SATMONN06	g218233	BLASTN	445	1e-51	86
1227	2462	700573201H1	SATMON030	g218233	BLASTN	691	1e-51	81
1228	2462	700452623H1	SATMON028	g218233	BLASTN	694	1e-51	85
1229	2462	700351523H1	SATMON023	g218233	BLASTN	679	1e-50	86
1230	2462	700042795H1	SATMON004	g218233	BLASTN	630	1e-45	86
1231	2462	700445979H1	SATMON027	g218233	BLASTN	595	1e-43	86
1232	2462	700201855H1	SATMON003	g218233	BLASTN	604	1e-43	87
1233	2462	700573101H1	SATMON030	g218233	BLASTN	594	1e-42	78
1234	2462	700049543H1	SATMON003	g218233	BLASTN	577	1e-41	79
1235	2462	700432359H1	SATMONN01	g218233	BLASTN	561	1e-40	81
1236	2462	701182021H1	SATMONN06	g218233	BLASTN	561	1e-40	85
1237	2462	701182019H1	SATMONN06	g218233	BLASTN	566	1e-40	86
1238	2462	700150928H1	SATMON007	g218233	BLASTN	569	1e-40	85
1239	2462	700202824H1	SATMON003	g218233	BLASTN	336	1e-39	86
1240	2462	700451056H1	SATMON028	g218233	BLASTN	553	1e-39	85
1241	2462	700449958H1	SATMON028	g218233	BLASTN	544	1e-38	86
1242	2462	700347592H1	SATMON023	g218233	BLASTN	403	1e-34	78
1243	2462	700573195H1	SATMON030	g218233	BLASTN	200	1e-22	84
1244	2462	700582836H1	SATMON031	g303849	BLASTX	157	1e-15	83
1245	2462	700029459H1	SATMON003	g303849	BLASTX	134	1e-11	84
1246	27065	700583429H1	SATMON031	g1064895	BLASTX	72	1e-13	54
1247	-L1482546	LIB148-007-Q1-E1-E6	LIB148	g218233	BLASTN	359	1e-19	75
1248	2462	LIB3067-039-Q1-K1-B10	LIB3067	g218233	BLASTN	711	1e-52	82
1249	2462	LIB3078-001-Q1-K1-F3	LIB3078	g218233	BLASTN	488	1e-49	85
1250	2462	LIB3067-029-Q1-K1-C3	LIB3067	g1236951	BLASTX	166	1e-31	96
1251	25174	LIB189-022-Q1-E1-E9	LIB189	g758643	BLASTN	440	1e-25	76

# MAIZE GLUCOSE-6-PHOSPHATE 1-DEHYDROGENASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
1252	-700047645	700047645H1	SATMON003	g471345	BLASTX	193	1e-21	58
1253	-700210379	700210379H1	SATMON016	g1480344	BLASTX	103	1e-10	85
1254	9135	700203121H1	SATMON003	g1166405	BLASTX	108	1e-10	78

# MAIZE PHOSPHOGLUCOMUTASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
1255	-700045655	700045655H1	SATMON004	g534982	BLASTX	144	1e-12	65
1256	-700053330	700053330H1	SATMON009	g3294467	BLASTX	211	1e-23	71
1257	-700102193	700102193H1	SATMON010	g534982	BLASTX	145	1e-14	53
1258	-700166982	700166982H1	SATMON013	g2795876	BLASTX	168	1e-16	52
1259	-700169540	700169540H1	SATMON013	g534982	BLASTX	180	1e-17	61
1260	-700210088	700210088H1	SATMON016	g534982	BLASTX	328	1e-38	55
1261	-700573194	700573194H1	SATMON030	g534982	BLASTX	192	1e-21	54
1262	-700616588	700616588H1	SATMON033	g3294468	BLASTN	593	1e-48	95
1263	119	700574655H1	SATMON030	g3294466	BLASTN	1705	1e-133	98
1264	119	700574672H1	SATMON030	g3294466	BLASTN	820	1e-121	100
1265	119	700100992H1	SATMON009	g3294466	BLASTN	1545	1e-119	99
1266	119	700615409H1	SATMON033	g3294466	BLASTN	1050	1e-118	100
1267	119	700210693H1	SATMON016	g3294468	BLASTN	1515	1e-117	100
1268	119	700381526H1	SATMON023	g3294468	BLASTN	1490	1e-115	100
1269	119	700026372H1	SATMON003	g3294466	BLASTN	1463	1e-113	99
1270	119	700201578H1	SATMON003	g3294468	BLASTN	677	1e-112	96
1271	119	700101083H1	SATMON009	g3294468	BLASTN	1430	1e-110	100
1272	119	700217101H1	SATMON016	g3294468	BLASTN	1420	1e-109	100
1273	119	700222466H1	SATMON011	g3294466	BLASTN	957	1e-106	97
1274	119	700072492H1	SATMON007	g3294466	BLASTN	1381	1e-106	99
1275	119	700043724H1	SATMON004	g3294468	BLASTN	1390	1e-106	100
1276	119	700346762H1	SATMON021	g3294468	BLASTN	1333	1e-102	94
1277	119	700347741H1	SATMON023	g3294468	BLASTN	1339	1e-102	97
1278	119	700550792H1	SATMON022	g3294466	BLASTN	731	1e-101	99
1279	119	700380144H1	SATMON021	g3294466	BLASTN	1216	1e-98	97
1280	119	700241526H1	SATMON010	g3294466	BLASTN	1285	1e-98	100
1281	119	700380456H1	SATMON021	g3294468	BLASTN	650	1e-97	99
1282	119	700238734H1	SATMON010	g3294466	BLASTN	974	1e-97	97
1283	119	700083634H1	SATMON011	g3294468	BLASTN	1265	1e-96	100
1284	119	700383086H1	SATMON024	g3294466	BLASTN	961	1e-94	96
1285	119	700169630H1	SATMON013	g3294466	BLASTN	1245	1e-94	100
1286	119	701177766H1	SATMONN05	g3294466	BLASTN	1187	1e-93	97
1287	119	700142461H1	SATMON012	g3294466	BLASTN	1231	1e-93	99
1288	119	700044235H1	SATMON004	g3294466	BLASTN	1175	1e-89	100
1289	119	700216921H1	SATMON016	g3294466	BLASTN	1165	1e-88	100
1290	119	700333779H1	SATMON019	g3294466	BLASTN	996	1e-87	96
1291	119	700021881H1	SATMON001	g3294468	BLASTN	1120	1e-84	100
1292	119	700049194H1	SATMON003	g3294468	BLASTN	940	1e-82	98
1293	119	700164477H1	SATMON013	g3294466	BLASTN	1091	1e-82	99
1294	119	700169514H1	SATMON013	g3294468	BLASTN	865	1e-80	100
1295	119	700050896H1	SATMON003	g3294466	BLASTN	591	1e-76	94
1296	119	700172394H1	SATMON013	g3294466	BLASTN	1024	1e-76	99
1297	119	700211437H1	SATMON016	g3294466	BLASTN	994	1e-73	99
1298	119	700084535H1	SATMON011	g3294468	BLASTN	973	1e-72	99
1299	119	700203439H1	SATMON003	g3294466	BLASTN	765	1e-71	100
1300	119	700257833H1	SATMON017	g3294468	BLASTN	611	1e-69	94
1301	119	700621831H1	SATMON034	g3294466	BLASTN	412	1e-52	90
1302	119	700354511H1	SATMON024	g3294468	BLASTN	703	1e-52	91
1303	119	700203525H1	SATMON003	g3294468	BLASTN	708	1e-50	99

1304	119	700020476H1	SATMON001	g3294468	BLASTN	658	1e-45	99
1305	119	700050562H1	SATMON003	g3294466	BLASTN	544	1e-42	88
1306	119	700613868H1	SATMON033	g3294466	BLASTN	615	1e-42	100
1307	119	700574982H1	SATMON030	g3294466	BLASTN	473	1e-35	97
1308	119	700049512H1	SATMON003	g3294466	BLASTN	268	1e-29	95
1309	119	700260372H2	SATMON017	g3294466	BLASTN	226	1e-10	89
1310	16726	700082801H1	SATMON011	g2829893	BLASTX	278	1e-30	55
1311	16726	700212054H1	SATMON016	g2829893	BLASTX	220	1e-23	53
1312	19462	700097450H1	SATMON009	g1814400	BLASTN	323	1e-29	64
1313	19462	700441165H1	SATMON026	g1408296	BLASTX	239	1e-25	61
1314	24348	700379424H1	SATMON020	g3294466	BLASTN	707	1e-50	98
1315	2587	700089556H1	SATMON011	g2829893	BLASTX	117	1e-8	67
1316	3016	700204345H1	SATMON003	g3294468	BLASTN	1784	1e-139	98
1317	3016	700098713H1	SATMON009	g3294468	BLASTN	1516	1e-117	99
1318	3016	700084751H1	SATMON011	g3294466	BLASTN	1475	1e-114	100
1319	3016	700351326H1	SATMON023	g3294468	BLASTN	1460	1e-112	100
1320	3016	700097161H1	SATMON009	g3294466	BLASTN	1308	1e-109	98
1321	3016	700266423H1	SATMON017	g3294468	BLASTN	1065	1e-108	96
1322	3016	700349605H1	SATMON023	g3294466	BLASTN	1335	1e-107	100
1323	3016	700350209H1	SATMON023	g3294468	BLASTN	1188	1e-106	97
1324	3016	700265291H1	SATMON017	g3294468	BLASTN	873	1e-100	98
1325	3016	700457572H1	SATMON029	g3294466	BLASTN	1288	1e-98	98
1326	3016	700334810H1	SATMON019	g3294468	BLASTN	863	1e-97	99
1327	3016	700194444H1	SATMON014	g3294466	BLASTN	1265	1e-96	100
1328	3016	700457426H1	SATMON029	g3294466	BLASTN	1236	1e-94	98
1329	3016	700210958H1	SATMON016	g3294466	BLASTN	1148	1e-92	98
1330	3016	700075135H1	SATMON007	g3294468	BLASTN	1219	1e-92	97
1331	3016	700152065H1	SATMON007	g3294466	BLASTN	1135	1e-90	99
1332	3016	700219672H1	SATMON011	g3294468	BLASTN	823	1e-89	99
1333	3016	700170425H1	SATMON013	g3294466	BLASTN	1110	1e-83	100
1334	3016	700153495H1	SATMON007	g3294468	BLASTN	640	1e-82	100
1335	3016	700348567H1	SATMON023	g3294468	BLASTN	557	1e-81	87
1336	3016	700803158H1	SATMON036	g3294468	BLASTN	630	1e-60	85
1337	3016	700264923H1	SATMON017	g3294468	BLASTN	340	1e-50	98
1338	3016	700615715H1	SATMON033	g3294466	BLASTN	567	1e-48	96
1339	3016	700027830H1	SATMON003	g3294468	BLASTN	632	1e-43	95
1340	3016	700350539H1	SATMON023	g3294466	BLASTN	333	1e-41	96
1341	4562	700044891H1	SATMON004	g3294466	BLASTN	650	1e-45	74
1342	4562	700215538H1	SATMON016	g3294466	BLASTN	555	1e-37	67
1343	9894	700220429H1	SATMON011	g3294468	BLASTN	1302	1e-99	99
1344	9894	700236461H1	SATMON010	g3294466	BLASTN	1054	1e-90	97
1345	-L30594453	LIB3059-042-Q1-K1-B5	LIB3059	g1814401	BLASTX	290	1e-49	58
1346	-L30605287	LIB3060-049-Q1-K1-B7	LIB3060	g534982	BLASTX	172	1e-34	77
1347	119	LIB3059-019-Q1-K1-H1	LIB3059	g1881692	BLASTN	2094	1e-165	98
1348	119	LIB3059-031-Q1-K1-H10	LIB3059	g1881692	BLASTN	1926	1e-151	96
1349	119	LIB3069-012-Q1-K1-F2	LIB3069	g1881692	BLASTN	1188	1e-146	90
1350	119	LIB36-019-Q1-E1-A7	LIB36	g1881692	BLASTN	1783	1e-139	90
1351	119	LIB3078-023-	LIB3078	g1881692	BLASTN	860	1e-124	87

1352	119	Q1-K1-C3 LIB3067-058-	LIB3067	g1881692	BLASTN	991	1e-114	99
1353	119	Q1-K1-G1 LIB3062-048-	LIB3062	g1881692	BLASTN	1181	1e-103	97
1354	119	Q1-K1-B7 LIB3069-023-	LIB3069	g1881692	BLASTN	1176	1e-87	84
1355	119	Q1-K1-G4 LIB3069-025-	LIB3069	g1881692	BLASTN	611	1e-65	91
1356	24348	Q1-K1-B6 LIB3066-043-	LIB3066	g1881692	BLASTN	560	1e-37	100
1357	24348	Q1-K1-F11 LIB3067-048-	LIB3067	g1881692	BLASTN	543	1e-36	99
1358	3016	Q1-K1-F3 LIB143-002-	LIB143	g2829893	BLASTX	224	1e-51	72
1359	3016	Q1-E1-C12 LIB189-034-	LIB189	g2829893	BLASTX	216	1e-48	68
1360	3016	Q1-E1-A11 LIB3069-043-	LIB3069	g1814401	BLASTX	98	1e-32	64
		Q1-K1-D5						

**MAIZE UDP-GLUCOSE PYROPHOSPHORYLASE**

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
1361	-700197315	700197315H1	SATMON014	g1388021	BLASTX	122	1e-9	70
1362	-700203530	700203530H1	SATMON003	g1212995	BLASTN	568	1e-38	78
1363	-700267284	700267284H1	SATMON017	g1212996	BLASTX	150	1e-13	87
1364	-700336683	700336683H1	SATMON019	g1752677	BLASTX	150	1e-27	82
1365	-700342324	700342324H1	SATMON021	g3107931	BLASTX	95	1e-14	80
1366	-700354856	700354856H1	SATMON024	g1388021	BLASTX	121	1e-22	75
1367	-700613858	700613858H1	SATMON033	g1212995	BLASTN	776	1e-59	88
1368	14982	700028996H1	SATMON003	g1212995	BLASTN	560	1e-37	76
1369	14982	700155115H1	SATMON007	g1212995	BLASTN	399	1e-31	81
1370	14982	700356747H1	SATMON024	g1388021	BLASTX	166	1e-15	76
1371	19537	700573761H1	SATMON030	g1212995	BLASTN	954	1e-70	79
1372	19537	700208049H1	SATMON016	g1212995	BLASTN	901	1e-66	78
1373	19537	700086382H1	SATMON011	g1212995	BLASTN	885	1e-64	77
1374	69	700091881H1	SATMON011	g1212995	BLASTN	844	1e-105	89
1375	69	700624406H1	SATMON034	g1212995	BLASTN	816	1e-97	88
1376	69	700211464H1	SATMON016	g1212995	BLASTN	1251	1e-95	88
1377	69	700099836H1	SATMON009	g1212995	BLASTN	1239	1e-94	88
1378	69	700084756H1	SATMON011	g1212995	BLASTN	1240	1e-94	90
1379	69	700076136H1	SATMON007	g1212995	BLASTN	1243	1e-94	89
1380	69	700073071H1	SATMON007	g1212995	BLASTN	1163	1e-88	86
1381	69	700614228H1	SATMON033	g1212995	BLASTN	1013	1e-87	84
1382	69	700379926H1	SATMON021	g1212995	BLASTN	1138	1e-86	88
1383	69	700089172H1	SATMON011	g1212995	BLASTN	1141	1e-86	88
1384	69	700265063H1	SATMON017	g1212995	BLASTN	1147	1e-86	86
1385	69	700085964H1	SATMON011	g1212995	BLASTN	1135	1e-85	85
1386	69	700282281H2	SATMON023	g1212995	BLASTN	1136	1e-85	86
1387	69	700429855H1	SATMONN01	g1212995	BLASTN	1114	1e-84	89
1388	69	700347453H1	SATMON023	g1212995	BLASTN	1117	1e-84	87
1389	69	700265087H1	SATMON017	g1212995	BLASTN	1120	1e-84	87
1390	69	700092705H1	SATMON008	g1212995	BLASTN	1122	1e-84	87
1391	69	700212686H1	SATMON016	g1212995	BLASTN	1123	1e-84	91

1392	69	700623332H1	SATMON034	g1212995	BLASTN	800	1e-83	86
1393	69	700041787H1	SATMON004	g1212995	BLASTN	1091	1e-82	91
1394	69	700219031H1	SATMON011	g1212995	BLASTN	1093	1e-82	89
1395	69	700218632H1	SATMON011	g1212995	BLASTN	1086	1e-81	90
1396	69	700211962H1	SATMON016	g1212995	BLASTN	1086	1e-81	86
1397	69	700220729H1	SATMON011	g1212995	BLASTN	916	1e-80	84
1398	69	700197025H1	SATMON014	g1212995	BLASTN	1063	1e-79	89
1399	69	700086546H1	SATMON011	g1212995	BLASTN	1049	1e-78	86
1400	69	700217064H1	SATMON016	g1212995	BLASTN	1051	1e-78	88
1401	69	700799128H1	SATMON036	g1212995	BLASTN	618	1e-77	88
1402	69	700265488H1	SATMON017	g1212995	BLASTN	1030	1e-77	84
1403	69	700043842H1	SATMON004	g1212995	BLASTN	1035	1e-77	87
1404	69	700236833H1	SATMON010	g1212995	BLASTN	1035	1e-77	87
1405	69	700219083H1	SATMON011	g1212995	BLASTN	1036	1e-77	88
1406	69	700042338H1	SATMON004	g1212995	BLASTN	1037	1e-77	87
1407	69	700352484H1	SATMON023	g1212995	BLASTN	1038	1e-77	85
1408	69	700083771H1	SATMON011	g1212995	BLASTN	613	1e-76	91
1409	69	700473855H1	SATMON025	g1212995	BLASTN	755	1e-76	85
1410	69	700353922H1	SATMON024	g1212995	BLASTN	1024	1e-76	85
1411	69	700023267H1	SATMON003	g1212995	BLASTN	1007	1e-75	89
1412	69	700157596H1	SATMON012	g1212995	BLASTN	1008	1e-75	87
1413	69	700218718H1	SATMON011	g1212995	BLASTN	1012	1e-75	86
1414	69	700162316H1	SATMON012	g1212995	BLASTN	626	1e-74	80
1415	69	700046475H1	SATMON004	g1212995	BLASTN	1003	1e-74	85
1416	69	700466010H1	SATMON025	g1212995	BLASTN	558	1e-73	82
1417	69	700571392H1	SATMON030	g1212995	BLASTN	985	1e-73	85
1418	69	700165241H1	SATMON013	g1212995	BLASTN	987	1e-73	85
1419	69	700457410H1	SATMON029	g1212995	BLASTN	988	1e-73	87
1420	69	700194672H1	SATMON014	g1212995	BLASTN	963	1e-71	86
1421	69	700089746H1	SATMON011	g1212995	BLASTN	964	1e-71	83
1422	69	700801620H1	SATMON036	g1212995	BLASTN	536	1e-70	91
1423	69	700264785H1	SATMON017	g1212995	BLASTN	952	1e-70	84
1424	69	700244093H1	SATMON010	g1212995	BLASTN	954	1e-70	85
1425	69	700043787H1	SATMON004	g1212995	BLASTN	957	1e-70	85
1426	69	700267269H1	SATMON017	g1212995	BLASTN	867	1e-69	85
1427	69	700167985H1	SATMON013	g1212995	BLASTN	940	1e-69	89
1428	69	700799042H1	SATMON036	g1212995	BLASTN	812	1e-68	89
1429	69	700163824H1	SATMON013	g1212995	BLASTN	888	1e-65	86
1430	69	700098307H1	SATMON009	g1212995	BLASTN	461	1e-63	81
1431	69	700805267H1	SATMON036	g1212995	BLASTN	734	1e-63	88
1432	69	700204843H1	SATMON003	g1212995	BLASTN	854	1e-62	88
1433	69	700206721H1	SATMON003	g1212995	BLASTN	461	1e-61	81
1434	69	700018559H1	SATMON001	g1212995	BLASTN	847	1e-61	85
1435	69	700026241H1	SATMON003	g1212995	BLASTN	847	1e-61	87
1436	69	700099987H1	SATMON009	g1212995	BLASTN	461	1e-60	81
1437	69	700475628H1	SATMON025	g1212995	BLASTN	750	1e-59	80
1438	69	700016675H1	SATMON001	g1212995	BLASTN	814	1e-59	86
1439	69	700150144H1	SATMON007	g1212995	BLASTN	822	1e-59	86
1440	69	700267260H1	SATMON017	g1212995	BLASTN	461	1e-58	80
1441	69	700261336H1	SATMON017	g1212995	BLASTN	564	1e-58	82
1442	69	700618652H1	SATMON033	g1212995	BLASTN	730	1e-58	78
1443	69	700469914H1	SATMON025	g1212995	BLASTN	735	1e-58	89
1444	69	700048027H1	SATMON003	g1212995	BLASTN	807	1e-58	83
1445	69	700165703H1	SATMON013	g1212995	BLASTN	796	1e-57	85

1446	69	700265403H1	SATMON017	g1212995	BLASTN	797	1e-57	78
1447	69	700099428H1	SATMON009	g1212995	BLASTN	474	1e-56	88
1448	69	700243212H1	SATMON010	g1212995	BLASTN	779	1e-56	84
1449	69	700092996H1	SATMON008	g1212995	BLASTN	789	1e-56	84
1450	69	700803035H1	SATMON036	g1212995	BLASTN	436	1e-54	80
1451	69	700235803H1	SATMON010	g1212995	BLASTN	688	1e-54	79
1452	69	700172581H1	SATMON013	g1212995	BLASTN	754	1e-54	79
1453	69	700214715H1	SATMON016	g1212995	BLASTN	762	1e-54	86
1454	69	700223082H1	SATMON011	g1212995	BLASTN	764	1e-54	84
1455	69	700093483H1	SATMON008	g1212995	BLASTN	357	1e-51	88
1456	69	700261920H1	SATMON017	g1212995	BLASTN	363	1e-51	82
1457	69	700221718H1	SATMON011	g1212995	BLASTN	363	1e-51	83
1458	69	700453106H1	SATMON028	g1212995	BLASTN	670	1e-51	82
1459	69	700210506H1	SATMON016	g1212995	BLASTN	461	1e-50	85
1460	69	700212333H1	SATMON016	g1212995	BLASTN	443	1e-49	83
1461	69	700072654H1	SATMON007	g1212995	BLASTN	443	1e-49	79
1462	69	700218282H1	SATMON016	g1212995	BLASTN	452	1e-49	85
1463	69	700263725H1	SATMON017	g1212995	BLASTN	662	1e-49	80
1464	69	700343083H1	SATMON021	g1212995	BLASTN	388	1e-48	80
1465	69	700219739H1	SATMON011	g1212995	BLASTN	443	1e-48	81
1466	69	700620336H1	SATMON034	g1212995	BLASTN	621	1e-48	88
1467	69	700264630H1	SATMON017	g1212995	BLASTN	377	1e-47	80
1468	69	700439242H1	SATMON026	g1212995	BLASTN	648	1e-47	83
1469	69	700259658H1	SATMON017	g1212995	BLASTN	511	1e-45	79
1470	69	700263521H1	SATMON017	g1212995	BLASTN	461	1e-44	79
1471	69	700261387H1	SATMON017	g1212995	BLASTN	461	1e-44	80
1472	69	700439277H1	SATMON026	g1212995	BLASTN	461	1e-43	84
1473	69	700452839H1	SATMON028	g1212995	BLASTN	544	1e-43	77
1474	69	700220236H1	SATMON011	g1212995	BLASTN	448	1e-40	84
1475	69	700472602H1	SATMON025	g1212995	BLASTN	254	1e-38	81
1476	69	700266424H1	SATMON017	g1212995	BLASTN	499	1e-37	80
1477	69	700449187H1	SATMON028	g1212995	BLASTN	540	1e-36	81
1478	69	700202731H1	SATMON003	g1212995	BLASTN	543	1e-36	79
1479	69	700156144H2	SATMON007	g1212995	BLASTN	441	1e-35	76
1480	69	700442679H1	SATMON026	g1212995	BLASTN	533	1e-35	80
1481	69	700449879H2	SATMON028	g1212995	BLASTN	535	1e-35	81
1482	69	700266832H1	SATMON017	g1212995	BLASTN	346	1e-34	77
1483	69	700332389H1	SATMON019	g1212995	BLASTN	382	1e-34	84
1484	69	700804202H1	SATMON036	g1212995	BLASTN	436	1e-34	76
1485	69	700151037H1	SATMON007	g1212995	BLASTN	443	1e-34	79
1486	69	700802810H1	SATMON036	g1212995	BLASTN	525	1e-34	85
1487	69	700455879H1	SATMON029	g1212995	BLASTN	448	1e-32	72
1488	69	700427769H1	SATMONN01	g1212995	BLASTN	481	1e-31	81
1489	69	700464626H1	SATMON025	g1212995	BLASTN	388	1e-30	76
1490	69	700439228H1	SATMON026	g1212995	BLASTN	470	1e-30	77
1491	69	700256847H1	SATMON017	g1212995	BLASTN	264	1e-29	85
1492	69	700204881H1	SATMON003	g1212995	BLASTN	430	1e-29	81
1493	69	700076032H1	SATMON007	g1212995	BLASTN	218	1e-26	72
1494	69	700426342H1	SATMONN01	g1212995	BLASTN	443	1e-26	79
1495	69	700209062H1	SATMON016	g1212995	BLASTN	279	1e-24	80
1496	69	700076988H1	SATMON007	g1212995	BLASTN	337	1e-24	83
1497	69	700349778H1	SATMON023	g1212995	BLASTN	406	1e-24	81
1498	69	700261886H1	SATMON017	g1212995	BLASTN	287	1e-15	80
1499	69	700426642H1	SATMONN01	g1388021	BLASTX	161	1e-14	76



1500	69	700155195H1	SATMON007	g1212995	BLASTN	155	1e-10	81
1501	69	700211992H1	SATMON016	g1212996	BLASTX	118	1e-9	85
1502	-L1485255	LIB148-053-Q1-E1-E12	LIB148	g1212995	BLASTN	691	1e-48	80
1503	-L30663959	LIB3066-015-Q1-K1-F12	LIB3066	g218000	BLASTN	251	1e-9	65
1504	19537	LIB3066-025-Q1-K1-E5	LIB3066	g1212995	BLASTN	1001	1e-74	79
1505	69	LIB3059-023-Q1-K1-C8	LIB3059	g1212995	BLASTN	1301	1e-133	89
1506	69	LIB3078-022-Q1-K1-C1	LIB3078	g1212995	BLASTN	1656	1e-129	86
1507	69	LIB3059-037-Q1-K1-H5	LIB3059	g1212995	BLASTN	1646	1e-128	86
1508	69	LIB3061-030-Q1-K1-A12	LIB3061	g1212995	BLASTN	1493	1e-124	86
1509	69	LIB3061-023-Q1-K1-A1	LIB3061	g1212995	BLASTN	1598	1e-124	86
1510	69	LIB3079-001-Q1-K1-D12	LIB3079	g1212995	BLASTN	1600	1e-124	83
1511	69	LIB189-028-Q1-E1-E3	LIB189	g1212995	BLASTN	1583	1e-123	87
1512	69	LIB3067-017-Q1-K1-D9	LIB3067	g1212995	BLASTN	1364	1e-120	88
1513	69	LIB3068-007-Q1-K1-F9	LIB3068	g1212995	BLASTN	1501	1e-116	85
1514	69	LIB3069-025-Q1-K1-E9	LIB3069	g1212995	BLASTN	1487	1e-115	85
1515	69	LIB3069-026-Q1-K1-E11	LIB3069	g1212995	BLASTN	1453	1e-112	85
1516	69	LIB3066-006-Q1-K1-G12	LIB3066	g1212995	BLASTN	1077	1e-107	83
1517	69	LIB3067-027-Q1-K1-D12	LIB3067	g1212995	BLASTN	1401	1e-107	86
1518	69	LIB189-010-Q1-E1-H10	LIB189	g1212995	BLASTN	1368	1e-105	85
1519	69	LIB3066-015-Q1-K1-G12	LIB3066	g1212995	BLASTN	1289	1e-104	82
1520	69	LIB3061-016-Q1-K1-G11	LIB3061	g1212995	BLASTN	1180	1e-102	84
1521	69	LIB3059-032-Q1-K1-G11	LIB3059	g1212995	BLASTN	1334	1e-102	87
1522	69	LIB3067-059-Q1-K1-G12	LIB3067	g1212995	BLASTN	1090	1e-100	85
1523	69	LIB3061-049-Q1-K1-C8	LIB3061	g1212995	BLASTN	1223	1e-98	79
1524	69	LIB3062-044-Q1-K1-F2	LIB3062	g1212995	BLASTN	1259	1e-96	83
1525	69	LIB3061-010-Q1-K1-F5	LIB3061	g1212995	BLASTN	1180	1e-95	84
1526	69	LIB3067-018-Q1-K1-A12	LIB3067	g1212995	BLASTN	1127	1e-89	82
1527	69	LIB3067-030-Q1-K1-F4	LIB3067	g1212995	BLASTN	1171	1e-88	83

1528	69	LIB3062-021-Q1-K1-F10	LIB3062	g1212995	BLASTN	1138	1e-86	87
1529	69	LIB3061-034-Q1-K1-D8	LIB3061	g1212995	BLASTN	1148	1e-86	85
1530	69	LIB3066-049-Q1-K1-C1	LIB3066	g1212995	BLASTN	1134	1e-85	83
1531	69	LIB3078-002-Q1-K1-F5	LIB3078	g1212995	BLASTN	859	1e-77	86
1532	69	LIB84-011-Q1-E1-G9	LIB84	g1212995	BLASTN	1020	1e-76	83
1533	69	LIB3067-043-Q1-K1-D2	LIB3067	g1212995	BLASTN	574	1e-59	77
1534	69	LIB189-003-Q1-E1-G5	LIB189	g1212995	BLASTN	247	1e-40	77
1535	69	LIB3062-008-Q1-K1-E6	LIB3062	g1212995	BLASTN	576	1e-37	63
1536	69	LIB189-016-Q1-E1-H7	LIB189	g1212996	BLASTX	156	1e-30	78
1537	69	LIB3067-007-Q1-K1-G4	LIB3067	g1212996	BLASTX	145	1e-28	82

#### SOYBEAN TRIOSE PHOSPHATE ISOMERASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
1538	-700743237	700743237H1	SOYMON012	g407525	BLASTX	173	1e-17	91
1539	-700977730	700977730H1	SOYMON009	g602589	BLASTN	373	1e-20	71
1540	-701056176	701056176H1	SOYMON032	g806311	BLASTN	752	1e-53	74
1541	-701110172	701110172H1	SOYMON036	g806311	BLASTN	801	1e-57	78
1542	10244	700995141H1	SOYMON011	g806311	BLASTN	470	1e-30	87
1543	10244	701124548H1	SOYMON037	g806311	BLASTN	490	1e-30	88
1544	10244	700739771H1	SOYMON012	g806311	BLASTN	329	1e-16	77
1545	10244	700999820H1	SOYMON018	g806312	BLASTX	147	84	
						1e-13		
1546	10244	701119858H1	SOYMON037	g806312	BLASTX	118	1e-9	72
1547	10535	700988684H1	SOYMON009	g806311	BLASTN	905	1e-66	79
1548	10535	700902425H1	SOYMON027	g806311	BLASTN	872	1e-63	80
1549	1357	701069004H1	SOYMON034	g806311	BLASTN	832	1e-60	81
1550	1357	701151554H1	SOYMON031	g806311	BLASTN	568	1e-38	82
1551	1357	700659936H1	SOYMON004	g806311	BLASTN	545	1e-36	79
1552	16	700680927H1	SOYMON008	g256119	BLASTN	1020	1e-81	78
1553	16	700656871H1	SOYMON004	g256119	BLASTN	903	1e-66	81
1554	16	701124364H1	SOYMON037	g256119	BLASTN	872	1e-64	80
1555	16	701134707H2	SOYMON038	g256119	BLASTN	874	1e-64	81
1556	16	700673750H1	SOYMON007	g256119	BLASTN	781	1e-60	81
1557	16	701123269H1	SOYMON037	g602589	BLASTN	819	1e-59	78
1558	16	701004846H1	SOYMON019	g256119	BLASTN	801	1e-58	80
1559	16	700993362H1	SOYMON011	g256119	BLASTN	808	1e-58	80
1560	16	701005445H1	SOYMON019	g256119	BLASTN	630	1e-56	78
1561	16	701134327H1	SOYMON038	g602589	BLASTN	782	1e-56	79
1562	16	701148169H1	SOYMON031	g602589	BLASTN	574	1e-51	76
1563	16	701153410H1	SOYMON031	g602589	BLASTN	451	1e-50	80
1564	16	700830168H1	SOYMON019	g256119	BLASTN	705	1e-50	77
1565	16	701120627H1	SOYMON037	g602589	BLASTN	715	1e-50	78

1566	16	700975358H1	SOYMON009	g602589	BLASTN	628	1e-49	77
1567	16	700755979H1	SOYMON014	g602589	BLASTN	697	1e-49	79
1568	16	701131374H1	SOYMON038	g602589	BLASTN	703	1e-49	79
1569	16	700994166H1	SOYMON011	g602589	BLASTN	513	1e-47	77
1570	16	701138038H1	SOYMON038	g602589	BLASTN	672	1e-47	77
1571	16	700974248H1	SOYMON005	g602589	BLASTN	658	1e-46	77
1572	16	700655832H1	SOYMON004	g602589	BLASTN	664	1e-46	78
1573	16	700758320H1	SOYMON015	g602589	BLASTN	409	1e-45	80
1574	16	701064709H1	SOYMON034	g602589	BLASTN	477	1e-45	78
1575	16	701138504H1	SOYMON038	g602589	BLASTN	591	1e-45	76
1576	16	700980284H1	SOYMON009	g602589	BLASTN	652	1e-45	79
1577	16	701133585H2	SOYMON038	g602589	BLASTN	634	1e-44	78
1578	16	700674706H1	SOYMON007	g602589	BLASTN	634	1e-44	78
1579	16	700964927H1	SOYMON022	g602589	BLASTN	639	1e-44	78
1580	16	700830923H1	SOYMON019	g602589	BLASTN	626	1e-43	76
1581	16	700662845H1	SOYMON005	g602589	BLASTN	617	1e-42	76
1582	16	701133824H1	SOYMON038	g602589	BLASTN	619	1e-42	78
1583	16	700848913H1	SOYMON021	g602589	BLASTN	603	1e-41	77
1584	16	701005984H1	SOYMON019	g602589	BLASTN	604	1e-41	78
1585	16	701140769H1	SOYMON038	g602589	BLASTN	605	1e-41	76
1586	16	700753357H1	SOYMON014	g602589	BLASTN	328	1e-40	78
1587	16	701056336H1	SOYMON032	g602589	BLASTN	344	1e-40	77
1588	16	700895411H1	SOYMON027	g602589	BLASTN	593	1e-40	78
1589	16	701060188H1	SOYMON033	g602589	BLASTN	277	1e-39	80
1590	16	700739461H1	SOYMON012	g602589	BLASTN	573	1e-39	77
1591	16	700941104H1	SOYMON024	g602589	BLASTN	579	1e-39	79
1592	16	700732960H1	SOYMON010	g602589	BLASTN	581	1e-39	78
1593	16	700686476H1	SOYMON008	g602589	BLASTN	583	1e-39	79
1594	16	701054231H1	SOYMON032	g602589	BLASTN	583	1e-39	77
1595	16	700671690H1	SOYMON006	g602589	BLASTN	566	1e-38	77
1596	16	700941174H1	SOYMON024	g602589	BLASTN	569	1e-38	78
1597	16	701125091H1	SOYMON037	g256119	BLASTN	358	1e-37	74
1598	16	700989827H1	SOYMON011	g602589	BLASTN	555	1e-37	78
1599	16	700835006H1	SOYMON019	g602589	BLASTN	555	1e-37	75
1600	16	700834847H1	SOYMON019	g602589	BLASTN	559	1e-37	78
1601	16	700953411H1	SOYMON022	g602589	BLASTN	314	1e-36	80
1602	16	700869222H1	SOYMON016	g602589	BLASTN	541	1e-36	78
1603	16	700850633H1	SOYMON023	g602589	BLASTN	544	1e-36	78
1604	16	700890283H1	SOYMON024	g602589	BLASTN	310	1e-35	80
1605	16	700727079H1	SOYMON009	g414549	BLASTN	358	1e-35	73
1606	16	700892544H1	SOYMON024	g602589	BLASTN	486	1e-35	78
1607	16	700869230H1	SOYMON016	g602589	BLASTN	528	1e-35	78
1608	16	700993034H1	SOYMON011	g602589	BLASTN	518	1e-34	75
1609	16	700975553H1	SOYMON009	g414549	BLASTN	524	1e-34	79
1610	16	700651326H1	SOYMON003	g602589	BLASTN	356	1e-33	80
1611	16	701215308H1	SOYMON035	g414549	BLASTN	450	1e-33	75
1612	16	700654480H1	SOYMON004	g414549	BLASTN	511	1e-33	80
1613	16	701045128H1	SOYMON032	g414549	BLASTN	512	1e-33	78
1614	16	701060759H1	SOYMON033	g414549	BLASTN	513	1e-33	80
1615	16	700741652H1	SOYMON012	g602589	BLASTN	493	1e-32	79
1616	16	700675469H1	SOYMON007	g602589	BLASTN	494	1e-32	78
1617	16	700657787H1	SOYMON004	g414549	BLASTN	495	1e-32	79
1618	16	701009957H2	SOYMON019	g414549	BLASTN	495	1e-32	80
1619	16	700983693H1	SOYMON009	g414549	BLASTN	495	1e-32	80

1620	16	701156784H1	SOYMON031	g602589	BLASTN	495	1e-32	78
1621	16	700893935H1	SOYMON024	g602589	BLASTN	481	1e-31	79
1622	16	701144619H1	SOYMON031	g414549	BLASTN	485	1e-31	78
1623	16	701148851H1	SOYMON031	g602589	BLASTN	487	1e-31	79
1624	16	701058218H1	SOYMON033	g602589	BLASTN	495	1e-31	78
1625	16	700975165H1	SOYMON009	g414549	BLASTN	466	1e-30	80
1626	16	701100165H1	SOYMON028	g602589	BLASTN	485	1e-30	79
1627	16	701150241H1	SOYMON031	g602589	BLASTN	455	1e-29	79
1628	16	701098308H1	SOYMON028	g414549	BLASTN	460	1e-29	79
1629	16	701150440H1	SOYMON031	g602589	BLASTN	462	1e-29	78
1630	16	700685125H1	SOYMON008	g414549	BLASTN	471	1e-29	81
1631	16	701061565H1	SOYMON033	g414549	BLASTN	471	1e-29	81
1632	16	700991418H1	SOYMON011	g602589	BLASTN	394	1e-28	68
1633	16	701156156H1	SOYMON031	g602589	BLASTN	456	1e-28	78
1634	16	701007231H2	SOYMON019	g602589	BLASTN	461	1e-28	79
1635	16	700829667H1	SOYMON019	g414549	BLASTN	333	1e-27	73
1636	16	701156033H1	SOYMON031	g602589	BLASTN	432	1e-27	78
1637	16	701014293H1	SOYMON019	g414549	BLASTN	446	1e-27	77
1638	16	701152138H1	SOYMON031	g414549	BLASTN	450	1e-27	81
1639	16	700945665H1	SOYMON024	g414549	BLASTN	450	1e-27	81
1640	16	701001407H1	SOYMON018	g169820	BLASTN	219	1e-26	72
1641	16	700983185H1	SOYMON009	g414549	BLASTN	435	1e-26	72
1642	16	700752364H1	SOYMON014	g414549	BLASTN	441	1e-26	76
1643	16	700992409H1	SOYMON011	g414549	BLASTN	427	1e-25	75
1644	16	701109396H1	SOYMON036	g414549	BLASTN	420	1e-24	76
1645	16	701151402H1	SOYMON031	g556171	BLASTX	151	1e-23	85
1646	16	701149617H1	SOYMON031	g556171	BLASTX	158	1e-23	86
1647	16	700747310H1	SOYMON013	g414549	BLASTN	406	1e-23	73
1648	16	701139569H1	SOYMON038	g556171	BLASTX	191	1e-22	84
1649	16	701213275H1	SOYMON035	g602589	BLASTN	255	1e-22	80
1650	16	701157185H1	SOYMON031	g556171	BLASTX	197	1e-20	90
1651	16	700655520H1	SOYMON004	g556171	BLASTX	166	1e-19	86
1652	16	701010779H1	SOYMON019	g556171	BLASTX	173	1e-19	64
1653	16	701044104H1	SOYMON032	g556171	BLASTX	188	1e-19	89
1654	16	700867605H1	SOYMON016	g556171	BLASTX	160	1e-17	70
1655	16	701058593H1	SOYMON033	g168647	BLASTX	169	1e-16	94
1656	16	701070286H1	SOYMON034	g168647	BLASTX	164	1e-15	91
1657	16	700877219H1	SOYMON018	g168647	BLASTX	154	1e-14	93
1658	16	700876790H1	SOYMON018	g168647	BLASTX	154	1e-14	93
1659	16	700877212H1	SOYMON018	g168647	BLASTX	154	1e-14	93
1660	16	700760847H1	SOYMON015	g556171	BLASTX	138	1e-13	86
1661	16	700893711H1	SOYMON024	g168647	BLASTX	140	1e-13	82
1662	16	700557532H1	SOYMON001	g256120	BLASTX	115	1e-12	88
1663	16	700793802H1	SOYMON017	g556171	BLASTX	138	1e-12	93
1664	16	700659725H1	SOYMON004	g556171	BLASTX	144	1e-12	47
1665	16	701044545H1	SOYMON032	g556171	BLASTX	144	1e-12	92
1666	16	701037485H1	SOYMON029	g556171	BLASTX	135	1e-11	96
1667	16	700683524H1	SOYMON008	g168647	BLASTX	136	1e-11	90
1668	16	700876711H1	SOYMON018	g168647	BLASTX	109	1e-10	85
1669	16	701155437H1	SOYMON031	g556171	BLASTX	130	1e-10	92
1670	28599	700997892H1	SOYMON018	g806311	BLASTN	834	1e-60	78
1671	31	701053174H1	SOYMON032	g806311	BLASTN	572	1e-37	73
1672	31	700754467H1	SOYMON014	g806312	BLASTX	145	1e-21	66
1673	31	701107430H1	SOYMON036	g806312	BLASTX	199	1e-20	63

1674	31	700985855H1	SOYMON009	g806312	BLASTX	145	1e-18	64
1675	31	701038167H1	SOYMON029	g806312	BLASTX	179	1e-17	61
1676	31	700670393H1	SOYMON006	g806312	BLASTX	167	1e-16	78
1677	31	700559280H1	SOYMON001	g609262	BLASTX	164	1e-15	69
1678	31	700793048H1	SOYMON017	g806312	BLASTX	97	1e-12	60
1679	31	700993683H1	SOYMON011	g806312	BLASTX	103	1e-11	60
1680	31	700663233H1	SOYMON005	g806312	BLASTX	130	1e-11	56
1681	31	700908079H1	SOYMON022	g806312	BLASTX	103	1e-10	60
1682	31	701043447H1	SOYMON029	g609262	BLASTX	126	1e-10	84
1683	31	700740188H1	SOYMON012	g806312	BLASTX	103	1e-8	60
1684	7466	700742922H1	SOYMON012	g806311	BLASTN	435	1e-27	76
1685	7466	700606255H1	SOYMON008	g806312	BLASTX	117	1e-17	80
1686	16	LIB3053-005-Q1-N1-F9	LIB3053	g602589	BLASTN	1000	1e-74	77
1687	16	LIB3039-035-Q1-E1-C5	LIB3039	g602589	BLASTN	979	1e-72	78
1688	16	LIB3039-031-Q1-E1-A8	LIB3039	g256119	BLASTN	911	1e-71	80
1689	16	LIB3030-003-Q1-B1-C9	LIB3030	g602589	BLASTN	949	1e-70	78
1690	16	LIB3039-023-Q1-E1-H12	LIB3039	g602589	BLASTN	913	1e-67	78
1691	16	LIB3039-047-Q1-E1-D8	LIB3039	g602589	BLASTN	566	1e-65	75
1692	16	LIB3039-052-Q1-E1-D6	LIB3039	g602589	BLASTN	890	1e-65	77
1693	16	LIB3039-051-Q1-E1-A1	LIB3039	g602589	BLASTN	855	1e-62	78
1694	16	LIB3049-009-Q1-E1-G5	LIB3049	g602589	BLASTN	783	1e-56	78
1695	16	LIB3039-009-Q1-E1-C1	LIB3039	g602589	BLASTN	805	1e-56	78
1696	16	LIB3055-006-Q1-N1-H3	LIB3055	g256119	BLASTN	481	1e-54	78
1697	16	LIB3055-013-Q1-N1-C3	LIB3055	g256119	BLASTN	769	1e-54	79
1698	16	LIB3049-034-Q1-E1-A2	LIB3049	g602589	BLASTN	626	1e-51	76
1699	16	LIB3049-022-Q1-E1-F9	LIB3049	g602589	BLASTN	519	1e-43	78
1700	16	LIB3049-030-Q1-E1-C7	LIB3049	g602589	BLASTN	572	1e-38	77
1701	16	LIB3040-035-Q1-E1-C5	LIB3040	g556171	BLASTX	175	1e-33	82
1702	16	LIB3040-005-Q1-E1-H8	LIB3040	g169820	BLASTN	324	1e-33	76
1703	16	LIB3028-025-Q1-B1-D1	LIB3028	g602589	BLASTN	464	1e-33	78
1704	16	LIB3039-022-Q1-E1-D5	LIB3039	g602589	BLASTN	357	1e-32	73
1705	16	LIB3052-001-Q1-B1-C5	LIB3052	G414549	BLASTN	327	1e-29	73
1706	28599	LIB3039-047-Q1-E1-D9	LIB3039	G806311	BLASTN	1183	1e-94	81

1707	28599	LIB3039-048-Q1-E1-D12	LIB3039	G806311	BLASTN	1007	1e-92	81
------	-------	-----------------------	---------	---------	--------	------	-------	----

# SOYBEAN FRUCTOSE 1,6-BISPHOSPHATE ALDOLASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
1708	-700565253	700565253H1	SOYMON002	G3021337	BLASTN	352	1e-39	76
1709	-700865276	700865276H1	SOYMON016	G3021337	BLASTN	629	1e-43	76
1710	-700873022	700873022H1	SOYMON018	G3696	BLASTX	211	1e-26	70
1711	-700943855	700943855H1	SOYMON024	G20204	BLASTX	202	1e-20	86
1712	-700974965	700974965H1	SOYMON005	g3021337	BLASTN	259	1e-10	84
1713	-701039850	701039850H1	SOYMON029	g22632	BLASTN	408	1e-23	76
1714	-701206840	701206840H1	SOYMON035	g3021338	BLASTX	151	1e-13	83
1715	11792	700654881H1	SOYMON004	g20204	BLASTX	150	1e-13	76
1716	11792	700746016H1	SOYMON013	g3021337	BLASTN	284	1e-12	67
1717	12314	701037190H1	SOYMON029	g3021337	BLASTN	634	1e-44	78
1718	12314	701042664H1	SOYMON029	g3021338	BLASTX	197	1e-20	66
1719	16	700651596H1	SOYMON003	g3021337	BLASTN	1101	1e-83	86
1720	16	700750439H1	SOYMON013	g3021337	BLASTN	1078	1e-81	86
1721	16	700649475H1	SOYMON003	g3021337	BLASTN	1082	1e-81	84
1722	16	700652995H1	SOYMON003	g3021337	BLASTN	1084	1e-81	82
1723	16	700981967H1	SOYMON009	g3021337	BLASTN	1071	1e-80	85
1724	16	700863243H1	SOYMON023	g3021337	BLASTN	1044	1e-78	86
1725	16	700558625H1	SOYMON001	g3021337	BLASTN	1041	1e-77	84
1726	16	700564806H1	SOYMON002	g3021337	BLASTN	1021	1e-76	80
1727	16	700746368H1	SOYMON013	g3021337	BLASTN	897	1e-75	86
1728	16	700960290H1	SOYMON022	g3021337	BLASTN	1009	1e-75	87
1729	16	701055132H1	SOYMON032	g3021337	BLASTN	1011	1e-75	86
1730	16	701056109H1	SOYMON032	g3021337	BLASTN	1012	1e-75	84
1731	16	701119884H1	SOYMON037	g3021337	BLASTN	1014	1e-75	87
1732	16	700898149H1	SOYMON027	g3021337	BLASTN	1015	1e-75	86
1733	16	700661436H1	SOYMON005	g3021337	BLASTN	596	1e-74	83
1734	16	701042223H1	SOYMON029	g3021337	BLASTN	997	1e-74	84
1735	16	700676004H1	SOYMON007	g3021337	BLASTN	984	1e-73	85
1736	16	700747718H1	SOYMON013	g3021337	BLASTN	988	1e-73	87
1737	16	700751133H1	SOYMON014	g3021337	BLASTN	989	1e-73	86
1738	16	701215247H1	SOYMON035	g3021337	BLASTN	989	1e-73	84
1739	16	700652484H1	SOYMON003	g3021337	BLASTN	910	1e-72	85
1740	16	700869785H1	SOYMON016	g3021337	BLASTN	970	1e-72	87
1741	16	700981960H1	SOYMON009	g3021337	BLASTN	970	1e-72	87
1742	16	700969335H1	SOYMON005	g3021337	BLASTN	972	1e-72	82
1743	16	700854174H1	SOYMON023	g3021337	BLASTN	965	1e-71	84
1744	16	700761638H1	SOYMON015	g3021337	BLASTN	966	1e-71	86
1745	16	700984860H1	SOYMON009	g3021337	BLASTN	967	1e-71	84
1746	16	701005716H1	SOYMON019	g3021337	BLASTN	967	1e-71	83
1747	16	700941053H1	SOYMON024	g3021337	BLASTN	968	1e-71	86
1748	16	700561358H1	SOYMON002	g3021337	BLASTN	968	1e-71	82
1749	16	700564906H1	SOYMON002	g3021337	BLASTN	562	1e-70	82
1750	16	700833951H1	SOYMON019	g3021337	BLASTN	954	1e-70	88
1751	16	701117626H1	SOYMON037	g3021337	BLASTN	957	1e-70	85
1752	16	700729103H1	SOYMON009	g3021337	BLASTN	535	1e-69	86
1753	16	700670615H1	SOYMON006	g3021337	BLASTN	936	1e-69	83
1754	16	701053635H1	SOYMON032	g3021337	BLASTN	941	1e-69	84









1917	16	700669492H1	SOYMON006	g3021337	BLASTN	504	1e-39	83
1918	16	700975340H1	SOYMON009	g3021337	BLASTN	574	1e-39	81
1919	16	700753528H1	SOYMON014	g3021337	BLASTN	576	1e-39	81
1920	16	700665923H1	SOYMON005	g3021337	BLASTN	373	1e-35	84
1921	16	701038320H1	SOYMON029	g3021337	BLASTN	518	1e-34	84
1922	16	700755605H1	SOYMON014	g3021337	BLASTN	431	1e-33	81
1923	16	700890349H1	SOYMON024	g3021337	BLASTN	511	1e-33	88
1924	16	700669817H1	SOYMON006	g3021337	BLASTN	363	1e-31	87
1925	16	701097640H1	SOYMON028	g3021337	BLASTN	476	1e-30	67
1926	16	700562959H1	SOYMON002	g3021337	BLASTN	482	1e-30	81
1927	16	700852454H1	SOYMON023	g3021337	BLASTN	446	1e-28	77
1928	16	701121443H1	SOYMON037	g3021337	BLASTN	418	1e-24	84
1929	16	701118247H1	SOYMON037	g3021337	BLASTN	280	1e-18	85
1930	16	700665401H1	SOYMON005	g927505	BLASTX	172	1e-16	94
1931	16	700750038H1	SOYMON013	g3021338	BLASTX	162	1e-15	84
1932	16	700665414H1	SOYMON005	g3021337	BLASTN	273	1e-13	88
1933	16	700889072H1	SOYMON024	g3021338	BLASTX	136	1e-11	83
1934	16	700727964H1	SOYMON009	g927505	BLASTX	137	1e-11	86
1935	16	700680648H1	SOYMON008	g3021337	BLASTN	226	1e-10	73
1936	16	701044547H1	SOYMON032	g927505	BLASTX	91	1e-9	76
1937	16	700649174H1	SOYMON003	g3021338	BLASTX	126	1e-9	83
1938	16531	701120682H1	SOYMON037	g3021337	BLASTN	716	1e-50	77
1939	1701	700993909H1	SOYMON011	g22633	BLASTX	112	1e-31	78
1940	1701	700955490H1	SOYMON022	g22633	BLASTX	176	1e-25	70
1941	1701	700682081H1	SOYMON008	g22633	BLASTX	147	1e-20	68
1942	1701	700988843H1	SOYMON011	g22633	BLASTX	90	1e-14	67
1943	1701	700740531H1	SOYMON012	g22633	BLASTX	92	1e-12	64
1944	1701	700790059H2	SOYMON011	g22633	BLASTX	92	1e-12	67
1945	1701	700872670H1	SOYMON018	g169037	BLASTX	144	1e-12	90
1946	1701	700990591H1	SOYMON011	g22632	BLASTN	199	1e-11	68
1947	1701	700743120H1	SOYMON012	g22633	BLASTX	92	1e-9	68
1948	1701	700994931H1	SOYMON011	g22633	BLASTX	92	1e-8	64
1949	1938	700738074H1	SOYMON012	g927507	BLASTX	134	1e-11	90
1950	239	701126904H1	SOYMON037	g169037	BLASTX	231	1e-24	81
1951	239	700668532H1	SOYMON006	g169037	BLASTX	202	1e-20	83
1952	239	700943660H1	SOYMON024	g169037	BLASTX	180	1e-17	84
1953	239	701009915H2	SOYMON019	g169037	BLASTX	180	1e-17	84
1954	239	701100047H2	SOYMON028	g169037	BLASTX	160	1e-15	84
1955	239	700794458H1	SOYMON017	g22633	BLASTX	131	1e-10	58
1956	239	700738441H1	SOYMON012	g169037	BLASTX	118	1e-8	78
1957	3425	700984050H1	SOYMON009	g3021337	BLASTN	874	1e-64	80
1958	3425	701014509H1	SOYMON019	g3021337	BLASTN	520	1e-60	80
1959	3425	701138819H1	SOYMON038	g3021337	BLASTN	815	1e-59	80
1960	3425	700977309H1	SOYMON009	g3021337	BLASTN	809	1e-58	80
1961	3425	700984876H1	SOYMON009	g3021337	BLASTN	813	1e-58	80
1962	3425	701046151H1	SOYMON032	g3021337	BLASTN	730	1e-52	80
1963	3425	700889668H1	SOYMON024	g3021337	BLASTN	737	1e-52	81
1964	3425	700976571H1	SOYMON009	g3021337	BLASTN	737	1e-52	81
1965	3425	701045371H1	SOYMON032	g3021337	BLASTN	716	1e-50	79
1966	3425	700548283H1	SOYMON002	g3021337	BLASTN	700	1e-49	81
1967	3425	701103461H1	SOYMON028	g3021337	BLASTN	705	1e-49	81
1968	3425	700898446H1	SOYMON027	g3021337	BLASTN	686	1e-48	83
1969	3425	701006432H1	SOYMON019	g3021337	BLASTN	688	1e-48	83
1970	3425	701041476H1	SOYMON029	g3021337	BLASTN	693	1e-48	81

1971	3425	700568335H1	SOYMON002	g3021337	BLASTN	678	1e-47	82
1972	3425	701046312H1	SOYMON032	g3021337	BLASTN	650	1e-45	85
1973	3425	701050171H1	SOYMON032	g3021337	BLASTN	650	1e-45	85
1974	3425	700685063H1	SOYMON008	g3021337	BLASTN	643	1e-44	83
1975	3425	701010250H2	SOYMON019	g3021337	BLASTN	542	1e-36	86
1976	3425	700665454H1	SOYMON005	g3021337	BLASTN	520	1e-34	80
1977	3425	701043888H1	SOYMON032	g3021337	BLASTN	495	1e-32	85
1978	3425	700726806H1	SOYMON009	g3021337	BLASTN	213	1e-23	76
1979	491	700997879H1	SOYMON018	g22632	BLASTN	789	1e-56	77
1980	491	700646208H1	SOYMON012	g22632	BLASTN	733	1e-52	76
1981	491	700559796H1	SOYMON001	g22632	BLASTN	715	1e-50	76
1982	491	700789784H1	SOYMON011	g22632	BLASTN	664	1e-46	76
1983	491	700683122H1	SOYMON008	g22632	BLASTN	485	1e-41	86
1984	491	701105914H1	SOYMON036	g22632	BLASTN	504	1e-41	73
1985	491	700558789H1	SOYMON001	g22632	BLASTN	607	1e-41	74
1986	491	700873051H1	SOYMON018	g22632	BLASTN	608	1e-41	75
1987	491	700684010H1	SOYMON008	g22632	BLASTN	597	1e-40	75
1988	491	700786096H2	SOYMON011	g22632	BLASTN	576	1e-39	75
1989	491	700731865H1	SOYMON010	g22632	BLASTN	582	1e-39	75
1990	491	701108111H1	SOYMON036	g22632	BLASTN	467	1e-38	75
1991	491	700740887H1	SOYMON012	g22632	BLASTN	567	1e-38	74
1992	491	700559579H1	SOYMON001	g22632	BLASTN	572	1e-38	75
1993	491	700996104H1	SOYMON018	g22632	BLASTN	476	1e-37	76
1994	491	700682145H1	SOYMON008	g22632	BLASTN	542	1e-36	74
1995	491	700737263H1	SOYMON010	g22632	BLASTN	526	1e-35	74
1996	491	700547963H1	SOYMON001	g22632	BLASTN	527	1e-35	73
1997	491	700686296H1	SOYMON008	g22632	BLASTN	527	1e-35	73
1998	491	700646072H1	SOYMON011	g22632	BLASTN	537	1e-35	74
1999	491	701106662H1	SOYMON036	g22632	BLASTN	514	1e-34	74
2000	491	700684335H1	SOYMON008	g22632	BLASTN	516	1e-34	74
2001	491	701000609H1	SOYMON018	g22632	BLASTN	520	1e-34	74
2002	491	700685658H1	SOYMON008	g22632	BLASTN	520	1e-34	74
2003	491	700875532H1	SOYMON018	g22632	BLASTN	521	1e-34	73
2004	491	700730264H1	SOYMON009	g22632	BLASTN	502	1e-33	74
2005	491	700872948H1	SOYMON018	g22632	BLASTN	502	1e-33	74
2006	491	700685813H1	SOYMON008	g22632	BLASTN	502	1e-33	74
2007	491	701104554H1	SOYMON036	g22632	BLASTN	503	1e-33	74
2008	491	700960601H1	SOYMON022	g22632	BLASTN	503	1e-33	74
2009	491	700876633H1	SOYMON018	g22632	BLASTN	503	1e-33	74
2010	491	700739662H1	SOYMON012	g22632	BLASTN	504	1e-33	72
2011	491	700685904H1	SOYMON008	g22632	BLASTN	505	1e-33	72
2012	491	700995183H1	SOYMON011	g22632	BLASTN	513	1e-33	73
2013	491	700901996H1	SOYMON027	g22632	BLASTN	513	1e-33	74
2014	491	700727070H1	SOYMON009	g22632	BLASTN	490	1e-32	72
2015	491	700685790H1	SOYMON008	g22632	BLASTN	492	1e-32	74
2016	491	700998652H1	SOYMON018	g22632	BLASTN	494	1e-32	72
2017	491	700740465H1	SOYMON012	g22632	BLASTN	482	1e-31	74
2018	491	700682621H2	SOYMON008	g22632	BLASTN	484	1e-31	74
2019	491	700874316H1	SOYMON018	g22632	BLASTN	466	1e-30	73
2020	491	700686477H1	SOYMON008	g22632	BLASTN	473	1e-30	73
2021	491	700739979H1	SOYMON012	g22632	BLASTN	476	1e-30	74
2022	491	700739416H1	SOYMON012	g22632	BLASTN	476	1e-30	74
2023	491	700685976H1	SOYMON008	g22632	BLASTN	476	1e-30	74
2024	491	700739629H1	SOYMON012	g22632	BLASTN	486	1e-30	70

2025	491	700989163H1	SOYMON011	g22632	BLASTN	468	1e-29	72
2026	491	701000555H1	SOYMON018	g22632	BLASTN	477	1e-29	72
2027	491	700872702H1	SOYMON018	g22632	BLASTN	436	1e-28	72
2028	491	701000781H1	SOYMON018	g22632	BLASTN	460	1e-28	73
2029	491	700682760H1	SOYMON008	g22632	BLASTN	463	1e-28	72
2030	491	700740390H1	SOYMON012	g22632	BLASTN	440	1e-27	73
2031	491	700685346H1	SOYMON008	g22632	BLASTN	451	1e-27	72
2032	491	700557272H1	SOYMON001	g22632	BLASTN	250	1e-26	78
2033	491	700953343H1	SOYMON022	g22632	BLASTN	349	1e-26	74
2034	491	700741960H1	SOYMON012	g22632	BLASTN	430	1e-26	73
2035	491	700680247H2	SOYMON008	g22632	BLASTN	425	1e-25	67
2036	491	700680002H2	SOYMON008	g22632	BLASTN	241	1e-24	72
2037	491	700684827H1	SOYMON008	g22632	BLASTN	379	1e-24	74
2038	491	700956353H1	SOYMON022	g22632	BLASTN	410	1e-24	72
2039	491	700787513H1	SOYMON011	g22632	BLASTN	235	1e-22	72
2040	491	700725070H1	SOYMON009	g22632	BLASTN	241	1e-22	71
2041	491	700741111H1	SOYMON012	g22632	BLASTN	304	1e-22	73
2042	491	700985308H1	SOYMON009	g22632	BLASTN	241	1e-21	80
2043	491	700738230H1	SOYMON012	g22632	BLASTN	241	1e-21	72
2044	491	700991396H1	SOYMON011	g22632	BLASTN	350	1e-21	72
2045	491	700741276H1	SOYMON012	g22632	BLASTN	379	1e-21	71
2046	491	700740223H1	SOYMON012	g22632	BLASTN	241	1e-20	72
2047	491	700738808H1	SOYMON012	g22632	BLASTN	241	1e-20	72
2048	491	700997995H1	SOYMON018	g22632	BLASTN	241	1e-19	81
2049	491	700875139H1	SOYMON018	g22632	BLASTN	241	1e-19	71
2050	491	700989713H1	SOYMON011	g22632	BLASTN	241	1e-19	73
2051	491	700958366H1	SOYMON022	g22632	BLASTN	241	1e-18	71
2052	491	700683887H1	SOYMON008	g22632	BLASTN	344	1e-18	70
2053	491	700740788H1	SOYMON012	g22632	BLASTN	339	1e-17	70
2054	491	700743058H1	SOYMON012	g22632	BLASTN	205	1e-16	81
2055	491	700996423H1	SOYMON018	g22632	BLASTN	234	1e-16	80
2056	491	700686075H1	SOYMON008	g22632	BLASTN	241	1e-16	71
2057	491	700738811H1	SOYMON012	g22632	BLASTN	193	1e-15	72
2058	491	700998312H1	SOYMON018	g22632	BLASTN	234	1e-15	73
2059	491	700681825H1	SOYMON008	g22632	BLASTN	241	1e-15	81
2060	491	701109105H1	SOYMON036	g22632	BLASTN	290	1e-14	69
2061	491	701203741H2	SOYMON035	g22632	BLASTN	230	1e-13	78
2062	491	700740785H1	SOYMON012	g22632	BLASTN	287	1e-13	68
2063	491	700738486H1	SOYMON012	g22632	BLASTN	295	1e-13	64
2064	491	700739078H1	SOYMON012	g22632	BLASTN	178	1e-12	73
2065	491	701002287H1	SOYMON018	g22632	BLASTN	255	1e-12	74
2066	491	700742470H1	SOYMON012	g22632	BLASTN	278	1e-12	69
2067	491	700743421H1	SOYMON012	g22632	BLASTN	261	1e-11	71
2068	491	700744039H1	SOYMON012	g22632	BLASTN	265	1e-11	69
2069	491	700789444H2	SOYMON011	g22632	BLASTN	158	1e-10	87
2070	491	700741074H1	SOYMON012	g22632	BLASTN	178	1e-10	77
2071	491	700998877H1	SOYMON018	g22632	BLASTN	235	1e-10	72
2072	491	700740005H1	SOYMON012	g22633	BLASTX	75	1e-9	64
2073	491	700872703H1	SOYMON018	g169037	BLASTX	116	1e-9	83
2074	491	700743301H1	SOYMON012	g22632	BLASTN	241	1e-9	76
2075	491	700875039H1	SOYMON018	g22632	BLASTN	241	1e-9	72
2076	491	700742515H1	SOYMON012	g22632	BLASTN	241	1e-9	76
2077	491	700990557H1	SOYMON011	g22632	BLASTN	241	1e-9	76
2078	491	700743995H1	SOYMON012	g22632	BLASTN	241	1e-9	76

2079	491	700743495H1	SOYMON012	g22632	BLASTN	241	1e-9	76
2080	491	701001909H1	SOYMON018	g22632	BLASTN	241	1e-9	76
2081	491	701001445H1	SOYMON018	g169037	BLASTX	115	1e-8	92
2082	491	700554881H1	SOYMON001	g169037	BLASTX	116	1e-8	94
2083	491	700954194H1	SOYMON022	g169037	BLASTX	116	1e-8	94
2084	491	700996869H1	SOYMON018	g22632	BLASTN	230	1e-8	76
2085	491	700897820H1	SOYMON027	g22632	BLASTN	234	1e-8	74
2086	491	700742574H1	SOYMON012	g22632	BLASTN	234	1e-8	74
2087	491	700684738H1	SOYMON008	g22632	BLASTN	235	1e-8	75
2088	7368	700739343H1	SOYMON012	g927507	BLASTX	164	1e-15	88
2089	-GM32379	LIB3051-015-Q1-E1-B12	LIB3051	g3021337	BLASTN	260	1e-28	77
2090	-GM8265	LIB3039-048-Q1-E1-F11	LIB3039	g3021337	BLASTN	481	1e-29	65
2091	16	LIB3027-010-Q1-B1-B7	LIB3027	g3021337	BLASTN	1393	1e-107	82
2092	16	LIB3039-049-Q1-E1-B8	LIB3039	g3021337	BLASTN	1297	1e-99	83
2093	16	LIB3051-061-Q1-K1-E11	LIB3051	g3021337	BLASTN	1303	1e-99	84
2094	16	LIB3056-009-Q1-N1-A10	LIB3056	g3021337	BLASTN	1126	1e-96	84
2095	16	LIB3051-025-Q1-K1-E11	LIB3051	g3021337	BLASTN	1262	1e-96	83
2096	16	LIB3056-014-Q1-N1-E1	LIB3056	g3021337	BLASTN	1077	1e-94	81
2097	16	LIB3055-005-Q1-N1-A8	LIB3055	g3021337	BLASTN	1227	1e-93	84
2098	16	LIB3040-045-Q1-E1-A4	LIB3040	g3021337	BLASTN	1211	1e-92	83
2099	16	LIB3028-010-Q1-B1-G9	LIB3028	g3021337	BLASTN	1215	1e-92	83
2100	16	LIB3056-010-Q1-N1-G8	LIB3056	g3021337	BLASTN	1217	1e-92	84
2101	16	LIB3039-029-Q1-E1-A6	LIB3039	g3021337	BLASTN	1128	1e-85	85
2102	16	LIB3051-014-Q1-E1-D2	LIB3051	g3021337	BLASTN	716	1e-80	83
2103	16	LIB3030-010-Q1-B1-D7	LIB3030	g3021337	BLASTN	1052	1e-78	83
2104	16	LIB3051-094-Q1-K1-A9	LIB3051	g3021337	BLASTN	778	1e-74	83
2105	16	LIB3028-030-Q1-B1-C9	LIB3028	g3021337	BLASTN	953	1e-70	85
2106	16	LIB3052-004-Q1-N1-D8	LIB3052	g3021337	BLASTN	868	1e-63	82
2107	16	LIB3065-014-Q1-N1-A3	LIB3065	g3021337	BLASTN	540	1e-61	79
2108	16	LIB3050-019-Q1-K1-H1	LIB3050	g168420	BLASTX	223	1e-40	63
2109	16	LIB3051-062-Q1-K1-B5	LIB3051	g3021337	BLASTN	541	1e-38	79
2110	3425	LIB3051-067-Q1-K1-E7	LIB3051	g3021337	BLASTN	1082	1e-81	78

2111	3425	LIB3050-006-Q1-E1-G7	LIB3050	g3021337	BLASTN	752	1e-57	75
2112	491	LIB3028-011-Q1-B1-B9	LIB3028	g22632	BLASTN	911	1e-67	75
2113	491	LIB3028-011-Q1-B1-F2	LIB3028	g22632	BLASTN	886	1e-65	77

# SOYBEAN FRUCTOSE-1,6-BISPHOSPHATASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
2114	-700685384	700685384H1	SOYMON008	g21244	BLASTN	597	1e-49	80
2115	-700737915	700737915H1	SOYMON012	g515746	BLASTN	1316	1e-100	97
2116	-700741457	700741457H1	SOYMON012	g3041774	BLASTN	692	1e-58	80
2117	-700874831	700874831H1	SOYMON018	g515746	BLASTN	1295	1e-99	100
2118	-700996155	700996155H1	SOYMON018	g3041774	BLASTN	651	1e-45	83
2119	-700996632	700996632H1	SOYMON018	g515746	BLASTN	507	1e-51	90
2120	-700998027	700998027H1	SOYMON018	g515746	BLASTN	636	1e-65	94
2121	-701209548	701209548H1	SOYMON035	g3041774	BLASTN	642	1e-44	83
2122	10129	700870828H1	SOYMON018	g21244	BLASTN	827	1e-60	79
2123	10129	700741669H1	SOYMON012	g21244	BLASTN	657	1e-53	80
2124	10348	700555754H1	SOYMON001	g21244	BLASTN	466	1e-29	77
2125	10348	700991527H1	SOYMON011	g440591	BLASTX	169	1e-16	88
2126	13716	700898719H1	SOYMON027	g515746	BLASTN	1186	1e-90	97
2127	13716	700993540H1	SOYMON011	g515746	BLASTN	1179	1e-89	98
2128	13716	700909657H1	SOYMON022	g515746	BLASTN	568	1e-57	86
2129	1894	700555054H1	SOYMON001	g515746	BLASTN	1320	1e-101	100
2130	1894	700685264H1	SOYMON008	g515746	BLASTN	1323	1e-101	99
2131	1894	700558854H1	SOYMON001	g515746	BLASTN	695	1e-98	100
2132	1894	700554755H1	SOYMON001	g515746	BLASTN	767	1e-98	99
2133	1894	701000504H1	SOYMON018	g515746	BLASTN	626	1e-95	98
2134	1894	700738115H1	SOYMON012	g515746	BLASTN	1230	1e-93	100
2135	1894	700992933H1	SOYMON011	g515746	BLASTN	1074	1e-91	98
2136	1894	701107444H1	SOYMON036	g515746	BLASTN	1201	1e-91	99
2137	1894	700852823H1	SOYMON023	g515746	BLASTN	1041	1e-90	98
2138	1894	700733478H1	SOYMON010	g515746	BLASTN	1150	1e-90	97
2139	1894	701105185H1	SOYMON036	g515746	BLASTN	641	1e-87	89
2140	1894	700737830H1	SOYMON012	g515746	BLASTN	1060	1e-87	100
2141	1894	700685110H1	SOYMON008	g515746	BLASTN	597	1e-86	90
2142	1894	700968307H1	SOYMON036	g515746	BLASTN	1113	1e-84	97
2143	1894	700653014H1	SOYMON003	g515746	BLASTN	587	1e-82	90
2144	1894	700555504H1	SOYMON001	g515746	BLASTN	626	1e-81	88
2145	1894	700751540H1	SOYMON014	g515746	BLASTN	585	1e-77	91
2146	1894	700901976H1	SOYMON027	g515746	BLASTN	505	1e-73	87
2147	1894	700986496H1	SOYMON009	g515746	BLASTN	559	1e-73	90
2148	1894	700751580H1	SOYMON014	g515746	BLASTN	569	1e-72	89
2149	1894	700751532H1	SOYMON014	g515746	BLASTN	571	1e-72	90
2150	1894	700990937H1	SOYMON011	g515746	BLASTN	544	1e-71	88
2151	1894	700740789H1	SOYMON012	g515746	BLASTN	630	1e-69	100
2152	1894	700743994H1	SOYMON012	g515746	BLASTN	945	1e-69	100
2153	1894	700754374H1	SOYMON014	g515746	BLASTN	460	1e-62	91
2154	1894	701001295H1	SOYMON018	g515746	BLASTN	541	1e-62	97
2155	1894	701155952H1	SOYMON031	g515746	BLASTN	568	1e-51	83
2156	1894	700872212H1	SOYMON018	g515746	BLASTN	670	1e-47	100
2157	1894	700682196H1	SOYMON008	g515746	BLASTN	609	1e-41	98

2158	1894	700738779H1	SOYMON012	g515746	BLASTN	252	1e-16	82
2159	26568	700844816H1	SOYMON021	g21244	BLASTN	649	1e-45	78
2160	27512	701128049H1	SOYMON037	g440591	BLASTX	185	1e-18	87
2161	7128	700649846H1	SOYMON003	g440591	BLASTX	125	1e-15	81
2162	10348	LIB3030-010- Q1-B1-C7	LIB3030	g21244	BLASTN	476	1e-28	76

#### FRUCTOSE-6-PHOSPHATE,2-KINASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
2163	-700730441	700730441H1	SOYMON009	g3309583	BLASTX	179	1e-17	82
2164	-700953509	700953509H1	SOYMON022	g3170229	BLASTN	674	1e-47	75
2165	-700955121	700955121H1	SOYMON022	g3309582	BLASTN	303	1e-14	68
2166	-GM28972	LIB3050-012- Q1-E1-E9	LIB3050	g3170229	BLASTN	1073	1e-80	80

#### SOYBEAN PHOSPHOGLUCOISOMERASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
2167	-700568558	700568558H1	SOYMON002	g1369950	BLASTX	165	1e-15	80
2168	-700845275	700845275H1	SOYMON021	g1100771	BLASTX	124	1e-10	53
2169	-700960755	700960755H1	SOYMON022	g1100771	BLASTX	153	1e-14	52
2170	18663	700838363H1	SOYMON020	g1100771	BLASTX	215	1e-22	63
2171	18663	700838355H1	SOYMON020	g1100771	BLASTX	155	1e-14	81
2172	19355	700897450H1	SOYMON027	g1100771	BLASTX	273	1e-31	74
2173	19355	700744258H1	SOYMON013	g1100771	BLASTX	207	1e-29	69
2174	19355	701153832H1	SOYMON031	g1100771	BLASTX	226	1e-23	58
2175	20088	700856114H1	SOYMON023	g1100771	BLASTX	176	1e-33	75
2176	20088	700670380H1	SOYMON006	g1100771	BLASTX	207	1e-33	71
2177	20088	700788785H2	SOYMON011	g1100771	BLASTX	120	1e-32	74
2178	20088	700847659H1	SOYMON021	g1100771	BLASTX	192	1e-31	84
2179	20088	701136417H1	SOYMON038	g1100771	BLASTX	169	1e-27	66
2180	31255	701207622H1	SOYMON035	g1100771	BLASTX	168	1e-29	61
2181	20088	LIB3051-014- Q1-E1-G3	LIB3051	g1100771	BLASTX	400	1e-68	73
2182	31255	LIB3056-008- Q1-N1-G8	LIB3056	g1100771	BLASTX	188	1e-52	62

#### SOYBEAN VACUOLAR H+-TRANSLOCATING-PYROPHOSPHATASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
2183	-700660662	700660662H1	SOYMON004	g16347	BLASTN	540	1e-36	79
2184	-700793860	700793860H1	SOYMON017	g2706449	BLASTN	808	1e-58	78
2185	-700837007	700837007H1	SOYMON020	g16347	BLASTN	776	1e-55	78
2186	-700890647	700890647H1	SOYMON024	g790474	BLASTN	826	1e-60	81
2187	-700942978	700942978H1	SOYMON024	g790478	BLASTN	605	1e-63	82
2188	-700944280	700944280H1	SOYMON024	g790479	BLASTX	119	1e-10	76
2189	-700974544	700974544H1	SOYMON005	g1103711	BLASTN	854	1e-62	83
2190	-700984449	700984449H1	SOYMON009	g1103711	BLASTN	287	1e-12	71
2191	-700989248	700989248H1	SOYMON011	g534915	BLASTN	276	1e-14	67
2192	-701102931	701102931H1	SOYMON028	g2706449	BLASTN	438	1e-46	76
2193	-701106870	701106870H1	SOYMON036	g790478	BLASTN	623	1e-47	75
2194	-701122796	701122796H1	SOYMON037	g2258074	BLASTX	71	1e-15	73

2195	-701132123	701132123H1	SOYMON038	g790478	BLASTN	627	1e-43	81
2196	-701136557	701136557H1	SOYMON038	g16347	BLASTN	376	1e-33	77
2197	14021	700973215H1	SOYMON005	g2668745	BLASTN	435	1e-39	80
2198	14021	701109310H1	SOYMON036	g2668745	BLASTN	281	1e-25	83
2199	16	700891764H1	SOYMON024	g790479	BLASTX	172	1e-16	68
2200	19232	701061126H1	SOYMON033	g790474	BLASTN	935	1e-69	81
2201	19232	700962864H1	SOYMON022	g790474	BLASTN	874	1e-64	82
2202	20872	700754883H1	SOYMON014	g790478	BLASTN	824	1e-59	81
2203	20872	700971147H1	SOYMON005	g1103711	BLASTN	564	1e-54	79
2204	2813	700797861H1	SOYMON017	g16347	BLASTN	731	1e-52	79
2205	2813	700944850H1	SOYMON024	g2570500	BLASTN	738	1e-52	82
2206	2813	701056207H1	SOYMON032	g2570500	BLASTN	556	1e-46	80
2207	2813	700605115H2	SOYMON003	g2570500	BLASTN	478	1e-42	80
2208	2813	700897063H1	SOYMON027	g2570500	BLASTN	596	1e-40	80
2209	2813	700561829H1	SOYMON002	g2570500	BLASTN	570	1e-38	80
2210	2813	701204883H1	SOYMON035	g2668745	BLASTN	545	1e-36	77
2211	2813	700754984H1	SOYMON014	g2570500	BLASTN	527	1e-35	75
2212	2813	700854552H1	SOYMON023	g2570500	BLASTN	536	1e-35	79
2213	2813	700873337H1	SOYMON018	g2570500	BLASTN	505	1e-33	75
2214	2813	700873349H1	SOYMON018	g2570500	BLASTN	506	1e-33	75
2215	2813	700952403H1	SOYMON022	g2668745	BLASTN	499	1e-32	76
2216	2813	700846561H1	SOYMON021	g2570500	BLASTN	488	1e-31	75
2217	2813	700953987H1	SOYMON022	g2570500	BLASTN	461	1e-29	75
2218	2813	700568667H1	SOYMON002	g2570500	BLASTN	296	1e-24	79
2219	2813	700895231H1	SOYMON024	g2258074	BLASTX	207	1e-22	80
2220	2813	701101791H1	SOYMON028	g2668746	BLASTX	147	1e-13	77
2221	8040	701121224H1	SOYMON037	g534915	BLASTN	298	1e-14	77
2222	8040	700743066H1	SOYMON012	g2668746	BLASTX	140	1e-12	80
2223	8531	701005139H1	SOYMON019	g2258073	BLASTN	871	1e-63	79
2224	8531	701008308H1	SOYMON019	g534915	BLASTN	789	1e-57	76
2225	8531	700559054H1	SOYMON001	g2570500	BLASTN	790	1e-57	77
2226	8531	700942540H1	SOYMON024	g2706449	BLASTN	755	1e-54	80
2227	8531	700790983H1	SOYMON011	g2258073	BLASTN	431	1e-52	77
2228	8531	701007949H1	SOYMON019	g2570500	BLASTN	404	1e-41	70
2229	8531	701123827H1	SOYMON037	g534915	BLASTN	436	1e-26	75
2230	8531	701013616H1	SOYMON019	g534915	BLASTN	431	1e-25	78
2231	8531	700565624H1	SOYMON002	g2570501	BLASTX	169	1e-16	85
2232	8531	701121092H1	SOYMON037	g2570501	BLASTX	110	1e-15	60
2233	16	LIB3040-003-Q1-E1-F6	LIB3040	g633598	BLASTN	523	1e-51	74
2234	16	LIB3051-114-Q1-K1-G5	LIB3051	g790478	BLASTN	457	1e-48	79
2235	16	LIB3039-020-Q1-E1-A2	LIB3039	g790478	BLASTN	338	1e-30	74
2236	2813	LIB3028-026-Q1-B1-B7	LIB3028	g2570500	BLASTN	1029	1e-77	80
2237	8040	LIB3049-045-Q1-E1-C3	LIB3049	g2706449	BLASTN	752	1e-52	72
2238	8040	LIB3049-005-Q1-E1-A7	LIB3049	g2570501	BLASTX	154	1e-32	61
2239	8531	LIB3050-013-Q1-E1-G8	LIB3050	g2570500	BLASTN	748	1e-53	72
2240	8531	LIB3073-025-Q1-K1-D6	LIB3073	g534915	BLASTN	711	1e-49	78



2241	8531	LIB3050-012-Q1-E1-D1	LIB3050	g2258074	BLASTX	93	1e-31	74
------	------	----------------------	---------	----------	--------	----	-------	----

**SOYBEAN PYROPHOSPHATE-DEPENDENT FRUCTOSE-6-PHOSPHATE  
PHOSPHOTRANSFERASE**

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
2242	7899	701008645H1	SOYMON019	g169538	BLASTX	160	1e-15	83

**INVERTASES**

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
2243	-700653543	700653543H1	SOYMON003	g1160487	BLASTN	541	1e-55	84
2244	-700992760	700992760H1	SOYMON011	g550319	BLASTX	117	1e-12	49
2245	-701005703	701005703H1	SOYMON019	g861157	BLASTX	213	1e-22	46
2246	-701047324	701047324H1	SOYMON032	g1160487	BLASTN	647	1e-45	81
2247	-701130328	701130328H1	SOYMON037	g167551	BLASTX	215	1e-22	61
2248	20460	700658149H1	SOYMON004	g861157	BLASTX	198	1e-20	72
2249	20460	701041452H1	SOYMON029	g402740	BLASTX	105	1e-13	76
2250	-GM31611	LIB3051-002-Q1-E1-B9	LIB3051	g1160487	BLASTN	1033	1e-77	77
2251	-GM34282	LIB3051-025-Q1-K1-C4	LIB3051	g1160487	BLASTN	1069	1e-80	79
2252	-GM34976	LIB3051-031-Q1-K1-A9	LIB3051	g1160487	BLASTN	769	1e-66	80
2253	31949	LIB3051-093-Q1-K1-B1	LIB3051	g1160487	BLASTN	948	1e-92	77
2254	31949	LIB3051-054-Q1-K2-D11	LIB3051	g1160487	BLASTN	903	1e-90	82

**SOYBEAN SUCROSE SYNTHASE**

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
2255	-700565776	700565776H1	SOYMON002	g3169544	BLASTX	89	1e-8	64
2256	-700606005	700606005H2	SOYMON007	g2570066	BLASTN	1069	1e-80	89
2257	-700664186	700664186H1	SOYMON005	g2606080	BLASTN	426	1e-62	91
2258	-700668119	700668119H1	SOYMON006	g2570066	BLASTN	279	1e-14	83
2259	-700668348	700668348H1	SOYMON006	g2570066	BLASTN	693	1e-48	88
2260	-700671225	700671225H1	SOYMON006	g16525	BLASTN	617	1e-42	72
2261	-700673918	700673918H1	SOYMON007	g218332	BLASTN	152	1e-9	92
2262	-700726266	700726266H1	SOYMON009	g2606080	BLASTN	237	1e-21	79
2263	-700747171	700747171H1	SOYMON013	g2606080	BLASTN	735	1e-52	89
2264	-700747359	700747359H1	SOYMON013	g218332	BLASTN	447	1e-28	78
2265	-700787443	700787443H2	SOYMON011	g22485	BLASTN	1171	1e-95	98
2266	-700796035	700796035H1	SOYMON017	g2570066	BLASTN	1039	1e-77	90
2267	-700832792	700832792H1	SOYMON019	g2606080	BLASTN	444	1e-31	88
2268	-700836673	700836673H1	SOYMON020	g2570066	BLASTN	843	1e-61	85
2269	-700841855	700841855H1	SOYMON020	g2570066	BLASTN	425	1e-35	84
2270	-700851758	700851758H1	SOYMON023	g2570066	BLASTN	211	1e-15	91
2271	-700851991	700851991H1	SOYMON023	g2570066	BLASTN	768	1e-55	81
2272	-700852943	700852943H1	SOYMON023	g2606080	BLASTN	250	1e-13	85
2273	-700853396	700853396H1	SOYMON023	g2570067	BLASTX	145	1e-13	65

2274	-700872206	700872206H1	SOYMON018	g1488570	BLASTX	235	1e-25	64
2275	-700876641	700876641H1	SOYMON018	g2606080	BLASTN	410	1e-53	88
2276	-700890526	700890526H1	SOYMON024	g2606080	BLASTN	652	1e-60	83
2277	-700893784	700893784H1	SOYMON024	g3169543	BLASTN	217	1e-11	82
2278	-700909222	700909222H1	SOYMON022	g2570066	BLASTN	440	1e-44	72
2279	-700944438	700944438H1	SOYMON024	g3169543	BLASTN	669	1e-46	73
2280	-700945733	700945733H1	SOYMON024	g1488569	BLASTN	504	1e-33	66
2281	-700969926	700969926H1	SOYMON005	g2570066	BLASTN	674	1e-47	72
2282	-701001986	701001986H1	SOYMON018	g1146237	BLASTX	106	1e-9	45
2283	-701005687	701005687H1	SOYMON019	g2606080	BLASTN	591	1e-40	85
2284	-701012195	701012195H1	SOYMON019	g2606080	BLASTN	418	1e-46	77
2285	-701046403	701046403H1	SOYMON032	g2606080	BLASTN	574	1e-38	76
2286	-701058966	701058966H1	SOYMON033	g218332	BLASTN	529	1e-56	84
2287	-701150574	701150574H1	SOYMON031	g1041247	BLASTX	155	1e-14	74
2288	-701205210	701205210H1	SOYMON035	g218332	BLASTN	981	1e-72	85
2289	10445	700605276H2	SOYMON003	g2606080	BLASTN	860	1e-65	84
2290	10445	700832417H1	SOYMON019	g2606080	BLASTN	876	1e-64	82
2291	10445	700833214H1	SOYMON019	g2606080	BLASTN	740	1e-58	83
2292	10445	700832409H1	SOYMON019	g2606080	BLASTN	800	1e-57	84
2293	10445	701007169H1	SOYMON019	g2606080	BLASTN	691	1e-55	81
2294	10445	701005913H1	SOYMON019	g2606080	BLASTN	680	1e-52	83
2295	10445	701204549H2	SOYMON035	g2606080	BLASTN	732	1e-52	83
2296	10445	701208347H1	SOYMON035	g2606080	BLASTN	656	1e-49	83
2297	10445	700958980H1	SOYMON022	g2606080	BLASTN	670	1e-49	83
2298	10445	700988126H1	SOYMON009	g2606080	BLASTN	324	1e-47	78
2299	10445	700830464H1	SOYMON019	g2606080	BLASTN	347	1e-47	79
2300	10445	700763911H1	SOYMON019	g3169543	BLASTN	517	1e-47	75
2301	10445	700891996H1	SOYMON024	g2606080	BLASTN	667	1e-46	88
2302	10445	700725104H1	SOYMON009	g2606080	BLASTN	577	1e-45	81
2303	10445	701124001H1	SOYMON037	g2606080	BLASTN	648	1e-45	86
2304	10445	700833919H1	SOYMON019	g2606080	BLASTN	496	1e-41	79
2305	10445	701006692H1	SOYMON019	g2606080	BLASTN	536	1e-41	86
2306	10445	700905349H1	SOYMON022	g2606080	BLASTN	585	1e-39	75
2307	10445	701204596H2	SOYMON035	g2606080	BLASTN	521	1e-38	79
2308	10445	700958885H1	SOYMON022	g2606080	BLASTN	351	1e-36	81
2309	10445	701208390H1	SOYMON035	g2606080	BLASTN	259	1e-29	86
2310	10445	701003131H1	SOYMON019	g2606080	BLASTN	442	1e-26	76
2311	10445	701207712H1	SOYMON035	g2606080	BLASTN	260	1e-17	78
2312	10445	701215107H1	SOYMON035	g2606080	BLASTN	260	1e-14	88
2313	10445	700852649H1	SOYMON023	g2606080	BLASTN	254	1e-13	74
2314	11259	701063407H1	SOYMON033	g2570066	BLASTN	1100	1e-82	87
2315	11259	700674761H1	SOYMON007	g2570066	BLASTN	739	1e-71	86
2316	11259	700839148H1	SOYMON020	g2570066	BLASTN	919	1e-67	87
2317	11259	700674815H1	SOYMON007	g2570066	BLASTN	904	1e-66	87
2318	12890	701103318H1	SOYMON028	g2570066	BLASTN	1005	1e-74	86
2319	12890	700855911H1	SOYMON023	g2570066	BLASTN	569	1e-69	86
2320	12890	700850874H1	SOYMON023	g2570066	BLASTN	937	1e-69	90
2321	12890	700837552H1	SOYMON020	g2570066	BLASTN	888	1e-65	89
2322	14264	700677058H1	SOYMON007	g2606080	BLASTN	578	1e-39	99
2323	14264	700679301H1	SOYMON007	g2606080	BLASTN	325	1e-18	90
2324	14740	701214452H1	SOYMON035	g2570066	BLASTN	1072	1e-80	89
2325	14740	701044972H1	SOYMON032	g2570066	BLASTN	537	1e-43	87
2326	14740	701040560H1	SOYMON029	g2570066	BLASTN	302	1e-24	75
2327	14740	700793901H1	SOYMON017	g2570066	BLASTN	231	1e-14	84

2328	15394	701136903H1	SOYMON038	g2606080	BLASTN	936	1e-69	81
2329	15394	701004431H1	SOYMON019	g218332	BLASTN	942	1e-69	80
2330	15394	701006153H1	SOYMON019	g218332	BLASTN	920	1e-67	83
2331	15394	701138281H1	SOYMON038	g218332	BLASTN	485	1e-40	82
2332	15394	701209319H1	SOYMON035	g3169543	BLASTN	508	1e-33	81
2333	16344	700746372H1	SOYMON013	g2606080	BLASTN	471	1e-65	85
2334	16344	700945706H1	SOYMON024	g2606080	BLASTN	635	1e-65	84
2335	17781	700960671H1	SOYMON022	g2570066	BLASTN	966	1e-71	88
2336	17781	700838540H1	SOYMON020	g2570066	BLASTN	532	1e-62	83
2337	20151	700847184H1	SOYMON021	g2570066	BLASTN	762	1e-72	90
2338	20151	700831558H1	SOYMON019	g2570066	BLASTN	980	1e-72	89
2339	22196	701046171H1	SOYMON032	g2606080	BLASTN	1321	1e-101	99
2340	22196	701207390H1	SOYMON035	g2606080	BLASTN	1258	1e-95	98
2341	25275	701013025H1	SOYMON019	g2606080	BLASTN	1353	1e-103	98
2342	25275	700561738H1	SOYMON002	g2606080	BLASTN	953	1e-84	91
2343	25380	700667735H1	SOYMON006	g2570066	BLASTN	959	1e-71	87
2344	25380	701047629H1	SOYMON032	g2570066	BLASTN	774	1e-55	89
2345	26818	701047072H1	SOYMON032	g2606080	BLASTN	830	1e-60	87
2346	26818	700737511H1	SOYMON010	g3169543	BLASTN	607	1e-57	83
2347	31182	701098655H1	SOYMON028	g2570066	BLASTN	951	1e-70	85
2348	318	701052316H1	SOYMON032	g2606080	BLASTN	1555	1e-120	100
2349	318	701053115H1	SOYMON032	g2606080	BLASTN	1281	1e-111	96
2350	318	700983049H1	SOYMON009	g2606080	BLASTN	1438	1e-110	96
2351	318	701058416H1	SOYMON033	g2606080	BLASTN	1385	1e-106	100
2352	318	701013289H1	SOYMON019	g2606080	BLASTN	1374	1e-105	99
2353	318	701002784H2	SOYMON019	g2606080	BLASTN	1365	1e-104	100
2354	318	700868516H1	SOYMON016	g2606080	BLASTN	1195	1e-103	100
2355	318	700978851H1	SOYMON009	g2606080	BLASTN	1325	1e-101	98
2356	318	701204954H1	SOYMON035	g2606080	BLASTN	770	1e-100	100
2357	318	700889102H1	SOYMON024	g2606080	BLASTN	1048	1e-100	99
2358	318	701053120H1	SOYMON032	g218332	BLASTN	1109	1e-100	90
2359	318	700731734H1	SOYMON010	g2606080	BLASTN	1308	1e-100	97
2360	318	700972625H1	SOYMON005	g2606080	BLASTN	1120	1e-98	99
2361	318	701006566H1	SOYMON019	g2606080	BLASTN	983	1e-97	99
2362	318	700952789H1	SOYMON022	g2606080	BLASTN	1276	1e-97	97
2363	318	701141518H1	SOYMON038	g2606080	BLASTN	716	1e-96	99
2364	318	700653475H1	SOYMON003	g3169543	BLASTN	1262	1e-96	87
2365	318	700650832H1	SOYMON003	g2606080	BLASTN	643	1e-95	97
2366	318	700678981H1	SOYMON007	g2606080	BLASTN	1142	1e-95	96
2367	318	700890311H1	SOYMON024	g2606080	BLASTN	1200	1e-95	100
2368	318	700892212H1	SOYMON024	g2606080	BLASTN	1250	1e-95	97
2369	318	700943424H1	SOYMON024	g2606080	BLASTN	1251	1e-95	99
2370	318	700833982H1	SOYMON019	g2606080	BLASTN	1255	1e-95	100
2371	318	700834361H1	SOYMON019	g2606080	BLASTN	981	1e-94	99
2372	318	700746379H1	SOYMON013	g2606080	BLASTN	1108	1e-94	96
2373	318	700889648H1	SOYMON024	g2606080	BLASTN	1238	1e-94	99
2374	318	701054868H1	SOYMON032	g2606080	BLASTN	1243	1e-94	95
2375	318	700959914H1	SOYMON022	g2606080	BLASTN	1226	1e-93	96
2376	318	701011518H1	SOYMON019	g2606080	BLASTN	705	1e-92	99
2377	318	700734053H1	SOYMON010	g2606080	BLASTN	765	1e-92	100
2378	318	701005295H1	SOYMON019	g2606080	BLASTN	962	1e-92	93
2379	318	700945690H1	SOYMON024	g2606080	BLASTN	1054	1e-92	99
2380	318	701118196H1	SOYMON037	g2606080	BLASTN	1100	1e-92	95
2381	318	700673512H1	SOYMON007	g2606080	BLASTN	1211	1e-92	97



2436	318	700788482H1	SOYMON011	g2606080	BLASTN	1038	1e-77	89
2437	318	701055686H1	SOYMON032	g2606080	BLASTN	1039	1e-77	90
2438	318	701054768H1	SOYMON032	g2606080	BLASTN	786	1e-76	88
2439	318	700854891H1	SOYMON023	g2606080	BLASTN	1030	1e-76	93
2440	318	701215276H1	SOYMON035	g2606080	BLASTN	1030	1e-76	90
2441	318	700944860H1	SOYMON024	g2606080	BLASTN	887	1e-75	96
2442	318	701010957H1	SOYMON019	g2606080	BLASTN	1011	1e-75	87
2443	318	701007175H1	SOYMON019	g2606080	BLASTN	1013	1e-75	90
2444	318	700725567H1	SOYMON009	g2606080	BLASTN	1013	1e-75	93
2445	318	700904972H1	SOYMON022	g22037	BLASTN	1015	1e-75	89
2446	318	700747391H1	SOYMON013	g2606080	BLASTN	1017	1e-75	87
2447	318	700747523H1	SOYMON013	g22037	BLASTN	836	1e-74	86
2448	318	700561819H1	SOYMON002	g218332	BLASTN	999	1e-74	82
2449	318	700835961H1	SOYMON019	g218332	BLASTN	1006	1e-74	87
2450	318	700562318H1	SOYMON002	g2606080	BLASTN	986	1e-73	84
2451	318	700745092H1	SOYMON013	g2606080	BLASTN	987	1e-73	88
2452	318	700832618H1	SOYMON019	g2606080	BLASTN	975	1e-72	87
2453	318	700891092H1	SOYMON024	g2606080	BLASTN	982	1e-72	88
2454	318	701119264H1	SOYMON037	g2606080	BLASTN	690	1e-71	89
2455	318	700894436H1	SOYMON024	g2606080	BLASTN	901	1e-71	91
2456	318	700894532H1	SOYMON024	g22037	BLASTN	959	1e-71	89
2457	318	700891712H1	SOYMON024	g22037	BLASTN	960	1e-71	89
2458	318	700895985H1	SOYMON027	g2606080	BLASTN	964	1e-71	89
2459	318	701203243H1	SOYMON035	g2606080	BLASTN	969	1e-71	88
2460	318	700985945H1	SOYMON009	g218332	BLASTN	713	1e-70	90
2461	318	700984768H1	SOYMON009	g2606080	BLASTN	781	1e-69	84
2462	318	700675710H1	SOYMON007	g2606080	BLASTN	784	1e-69	91
2463	318	700829561H1	SOYMON019	g218332	BLASTN	935	1e-69	87
2464	318	700964918H1	SOYMON022	g22037	BLASTN	942	1e-69	83
2465	318	701046747H1	SOYMON032	g2606080	BLASTN	422	1e-68	84
2466	318	700745512H1	SOYMON013	g3169543	BLASTN	457	1e-68	85
2467	318	700666671H1	SOYMON005	g218332	BLASTN	506	1e-68	87
2468	318	700889555H1	SOYMON024	g3169543	BLASTN	930	1e-68	86
2469	318	701147844H1	SOYMON031	g3169543	BLASTN	932	1e-68	86
2470	318	701206247H1	SOYMON035	g3169543	BLASTN	934	1e-68	82
2471	318	701103801H1	SOYMON036	g218332	BLASTN	723	1e-67	88
2472	318	700943746H1	SOYMON024	g218332	BLASTN	913	1e-67	86
2473	318	700745956H1	SOYMON013	g22037	BLASTN	921	1e-67	83
2474	318	700893512H1	SOYMON024	g218332	BLASTN	835	1e-66	90
2475	318	700897675H1	SOYMON027	g22037	BLASTN	899	1e-66	83
2476	318	700565777H1	SOYMON002	g2606080	BLASTN	510	1e-65	89
2477	318	700749851H1	SOYMON013	g2606080	BLASTN	887	1e-65	89
2478	318	700746286H1	SOYMON013	g2606080	BLASTN	876	1e-64	82
2479	318	700869142H1	SOYMON016	g2606080	BLASTN	885	1e-64	100
2480	318	700892442H1	SOYMON024	g2606080	BLASTN	872	1e-63	84
2481	318	700964153H1	SOYMON022	g22037	BLASTN	873	1e-63	83
2482	318	700898176H1	SOYMON027	g3169543	BLASTN	873	1e-63	84
2483	318	701056245H1	SOYMON032	g218332	BLASTN	543	1e-61	84
2484	318	700835360H1	SOYMON019	g218332	BLASTN	839	1e-61	88
2485	318	700749067H1	SOYMON013	g3169543	BLASTN	473	1e-60	86
2486	318	701008962H1	SOYMON019	g3169543	BLASTN	614	1e-60	90
2487	318	700980315H1	SOYMON009	g3169543	BLASTN	655	1e-60	84
2488	318	701202680H1	SOYMON035	g2606080	BLASTN	678	1e-60	89
2489	318	701202364H1	SOYMON035	g2606080	BLASTN	711	1e-60	85

2490	318	701037195H1	SOYMON029	g218332	BLASTN	439	1e-59	86
2491	318	701011681H1	SOYMON019	g3169543	BLASTN	459	1e-59	83
2492	318	700976368H1	SOYMON009	g218332	BLASTN	363	1e-58	85
2493	318	700829847H1	SOYMON019	g218332	BLASTN	384	1e-58	86
2494	318	700561920H1	SOYMON002	g2606080	BLASTN	809	1e-58	88
2495	318	701004573H1	SOYMON019	g2606080	BLASTN	813	1e-58	77
2496	318	701049462H1	SOYMON032	g3169543	BLASTN	450	1e-57	82
2497	318	700866272H1	SOYMON016	g3169543	BLASTN	421	1e-54	77
2498	318	700892632H1	SOYMON024	g2606080	BLASTN	453	1e-54	84
2499	318	701215184H1	SOYMON035	g218332	BLASTN	464	1e-54	88
2500	318	700831177H1	SOYMON019	g2606080	BLASTN	759	1e-54	85
2501	318	700835115H1	SOYMON019	g2606080	BLASTN	762	1e-54	81
2502	318	701015056H1	SOYMON019	g3169543	BLASTN	447	1e-53	81
2503	318	700675496H1	SOYMON007	g2606080	BLASTN	465	1e-53	95
2504	318	701052767H1	SOYMON032	g2606080	BLASTN	753	1e-53	88
2505	318	700833078H1	SOYMON019	g3169543	BLASTN	414	1e-52	84
2506	318	700869165H1	SOYMON016	g3169543	BLASTN	534	1e-51	84
2507	318	700831532H1	SOYMON019	g2606080	BLASTN	655	1e-51	100
2508	318	701010104H2	SOYMON019	g2606080	BLASTN	698	1e-51	85
2509	318	700890513H1	SOYMON024	g22037	BLASTN	575	1e-50	88
2510	318	700890952H1	SOYMON024	g2606080	BLASTN	709	1e-50	75
2511	318	700567301H1	SOYMON002	g22037	BLASTN	716	1e-50	82
2512	318	700945284H1	SOYMON024	g3169543	BLASTN	701	1e-49	75
2513	318	701206626H1	SOYMON035	g3169543	BLASTN	702	1e-49	81
2514	318	700748456H1	SOYMON013	g2606080	BLASTN	384	1e-48	77
2515	318	700981883H1	SOYMON009	g2606080	BLASTN	419	1e-48	85
2516	318	700942575H1	SOYMON024	g22037	BLASTN	340	1e-46	82
2517	318	700945125H1	SOYMON024	g2606080	BLASTN	405	1e-46	81
2518	318	700830469H1	SOYMON019	g3169543	BLASTN	636	1e-44	83
2519	318	700991669H1	SOYMON011	g218332	BLASTN	630	1e-43	83
2520	318	700866064H1	SOYMON016	g3169543	BLASTN	453	1e-41	84
2521	318	700866806H1	SOYMON016	g218332	BLASTN	607	1e-41	96
2522	318	700893154H1	SOYMON024	g2606080	BLASTN	539	1e-38	87
2523	318	700893118H1	SOYMON024	g2606080	BLASTN	539	1e-38	87
2524	318	701142963H2	SOYMON038	g218332	BLASTN	569	1e-38	90
2525	318	700945968H1	SOYMON024	g218332	BLASTN	572	1e-38	86
2526	318	700945788H1	SOYMON024	g2606080	BLASTN	514	1e-36	90
2527	318	700563455H1	SOYMON002	g2606080	BLASTN	496	1e-32	83
2528	318	700888936H1	SOYMON024	g3169543	BLASTN	498	1e-32	86
2529	318	701039594H1	SOYMON029	g22037	BLASTN	254	1e-28	84
2530	318	701015024H1	SOYMON019	g218333	BLASTX	65	1e-14	66
2531	318	700893166H1	SOYMON024	g22037	BLASTN	232	1e-8	85
2532	4258	700646449H1	SOYMON013	g22037	BLASTN	584	1e-39	70
2533	4258	700952838H1	SOYMON022	g20373	BLASTN	557	1e-37	70
2534	4413	700902256H1	SOYMON027	g2606080	BLASTN	1215	1e-99	97
2535	4413	700900032H1	SOYMON027	g2606080	BLASTN	720	1e-95	98
2536	4413	701006182H1	SOYMON019	g2606080	BLASTN	1179	1e-89	99
2537	4413	700831710H1	SOYMON019	g2606080	BLASTN	1070	1e-80	97
2538	4413	701008850H1	SOYMON019	g2606080	BLASTN	999	1e-74	99
2539	4413	701015432H1	SOYMON019	g2606080	BLASTN	813	1e-68	95
2540	4413	700987094H1	SOYMON009	g2606080	BLASTN	928	1e-68	84
2541	4413	700736179H1	SOYMON010	g2606080	BLASTN	753	1e-63	96
2542	4413	700890230H1	SOYMON024	g2606080	BLASTN	798	1e-57	95
2543	4413	701015314H1	SOYMON019	g2606080	BLASTN	639	1e-49	97

2544	4413	701052019H1	SOYMON032	g2606080	BLASTN	448	1e-37	95
2545	4748	701209527H1	SOYMON035	g2606080	BLASTN	1207	1e-91	93
2546	4748	700561984H1	SOYMON002	g2606080	BLASTN	542	1e-81	94
2547	4748	700895166H1	SOYMON024	g2606080	BLASTN	1004	1e-74	98
2548	4748	700843735H1	SOYMON021	g2606080	BLASTN	227	1e-20	93
2549	869	700650545H1	SOYMON003	g2606080	BLASTN	804	1e-107	94
2550	869	701205255H1	SOYMON035	g2606080	BLASTN	1135	1e-101	98
2551	869	700562091H1	SOYMON002	g2606080	BLASTN	1311	1e-100	92
2552	869	701213906H1	SOYMON035	g2606080	BLASTN	1300	1e-99	100
2553	869	700567712H1	SOYMON002	g2606080	BLASTN	634	1e-95	97
2554	869	701010943H1	SOYMON019	g2606080	BLASTN	1236	1e-94	99
2555	869	701006976H1	SOYMON019	g2606080	BLASTN	601	1e-93	98
2556	869	700752409H1	SOYMON014	g2606080	BLASTN	1080	1e-92	100
2557	869	701204769H1	SOYMON035	g2606080	BLASTN	795	1e-90	100
2558	869	701042737H1	SOYMON029	g2606080	BLASTN	1058	1e-90	99
2559	869	700832091H1	SOYMON019	g2606080	BLASTN	1116	1e-88	99
2560	869	701049161H1	SOYMON032	g2606080	BLASTN	1053	1e-86	96
2561	869	700906541H1	SOYMON022	g2606080	BLASTN	1087	1e-86	96
2562	869	701008182H1	SOYMON019	g2606080	BLASTN	1111	1e-86	92
2563	869	700831609H1	SOYMON019	g2606080	BLASTN	611	1e-84	92
2564	869	700834954H1	SOYMON019	g2606080	BLASTN	835	1e-84	100
2565	869	701037284H1	SOYMON029	g2606080	BLASTN	858	1e-83	94
2566	869	700561458H1	SOYMON002	g2606080	BLASTN	1019	1e-83	93
2567	869	701208357H1	SOYMON035	g2606080	BLASTN	1113	1e-83	99
2568	869	700747138H1	SOYMON013	g2606080	BLASTN	985	1e-80	93
2569	869	701014835H1	SOYMON019	g2606080	BLASTN	891	1e-78	89
2570	869	700956359H1	SOYMON022	g2606080	BLASTN	1052	1e-78	96
2571	869	701012740H1	SOYMON019	g2606080	BLASTN	643	1e-77	93
2572	869	701042523H1	SOYMON029	g2606080	BLASTN	667	1e-74	95
2573	869	701205775H1	SOYMON035	g2606080	BLASTN	745	1e-74	100
2574	869	701049184H1	SOYMON032	g2606080	BLASTN	600	1e-72	95
2575	869	700889179H1	SOYMON024	g2606080	BLASTN	942	1e-69	92
2576	869	700963920H1	SOYMON022	g2606080	BLASTN	718	1e-66	90
2577	869	700737476H1	SOYMON010	g2606080	BLASTN	548	1e-44	97
2578	869	701044544H1	SOYMON032	g2606080	BLASTN	462	1e-43	96
2579	869	700737636H1	SOYMON010	g2606080	BLASTN	426	1e-34	95
2580	9398	700837013H1	SOYMON020	g2570066	BLASTN	1025	1e-76	88
2581	9398	700891526H1	SOYMON024	g2570066	BLASTN	868	1e-63	87
2582	14740	LIB3051-038-Q1-K1-E10	LIB3051	g2570066	BLASTN	1331	1e-102	86
2583	31182	LIB3051-015-Q1-E1-F1	LIB3051	g2570066	BLASTN	1540	1e-119	88
2584	318	LIB3050-024-Q1-K1-H5	LIB3050	g2606080	BLASTN	1736	1e-135	95
2585	318	LIB3050-012-Q1-E1-F10	LIB3050	g2606080	BLASTN	1564	1e-125	98
2586	318	LIB3056-013-Q1-N1-H11	LIB3056	g3169543	BLASTN	1617	1e-125	86
2587	318	LIB3028-026-Q1-B1-F6	LIB3028	g3169543	BLASTN	1393	1e-107	84
2588	318	LIB3049-031-Q1-E1-B6	LIB3049	g3169543	BLASTN	1290	1e-98	90
2589	33428	LIB3051-085-Q1-K1-D11	LIB3051	g2570066	BLASTN	679	1e-53	86

2590 869 LIB3056-014- Q1-N1-G8 LIB3056 g2606080 BLASTN 1503 1e-132 96

SOYBEAN HEXOKINASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
2591	-700560085	700560085H1	SOYMON001	g1899024	BLASTN	456	1e-27	67
2592	-700752579	700752579H1	SOYMON014	g836808	BLASTX	113	1e-8	54
2593	-700753182	700753182H1	SOYMON014	g619928	BLASTX	234	1e-25	63
2594	-700838622	700838622H1	SOYMON020	g619927	BLASTN	767	1e-55	78
2595	-700840271	700840271H1	SOYMON020	g619927	BLASTN	525	1e-34	67
2596	-700844132	700844132H1	SOYMON021	g619927	BLASTN	474	1e-51	77
2597	-700898308	700898308H1	SOYMON027	g619927	BLASTN	464	1e-29	72
2598	-700904279	700904279H1	SOYMON022	g881521	BLASTX	129	1e-10	67
2599	-700904320	700904320H1	SOYMON022	g1899024	BLASTN	612	1e-42	71
2600	-700946357	700946357H1	SOYMON024	g619928	BLASTX	112	1e-18	69
2601	-700998007	700998007H1	SOYMON018	g1899024	BLASTN	367	1e-20	71
2602	-701097096	701097096H1	SOYMON028	g619927	BLASTN	488	1e-30	73
2603	-701102877	701102877H1	SOYMON028	g619927	BLASTN	551	1e-37	70
2604	-701103285	701103285H1	SOYMON028	g619928	BLASTX	179	1e-17	77
2605	-701105838	701105838H1	SOYMON036	g619928	BLASTX	274	1e-30	63
2606	-701138291	701138291H1	SOYMON038	g619927	BLASTN	819	1e-59	79
2607	12404	701065794H1	SOYMON034	g3087888	BLASTX	84	1e-11	44
2608	12404	701131030H1	SOYMON038	g1899025	BLASTX	120	1e-9	45
2609	12693	700846513H1	SOYMON021	g619927	BLASTN	459	1e-28	70
2610	12693	700656744H1	SOYMON004	g619927	BLASTN	251	1e-10	57
2611	12917	700906858H1	SOYMON022	g3087888	BLASTX	183	1e-32	80
2612	12917	700830011H1	SOYMON019	g619927	BLASTN	495	1e-32	70
2613	12917	701068501H1	SOYMON034	g619927	BLASTN	475	1e-29	72
2614	12917	701153981H1	SOYMON031	g3087887	BLASTN	440	1e-26	69
2615	222	700663332H1	SOYMON005	g619927	BLASTN	724	1e-51	76
2616	222	701142003H1	SOYMON038	g881520	BLASTN	542	1e-39	72
2617	222	700657213H1	SOYMON004	g881520	BLASTN	524	1e-34	73
2618	222	700833679H1	SOYMON019	g1899024	BLASTN	453	1e-28	80
2619	222	700556060H1	SOYMON001	g619927	BLASTN	463	1e-28	82
2620	23610	700984359H1	SOYMON009	g1899024	BLASTN	611	1e-42	73
2621	23610	701003284H1	SOYMON019	g1899024	BLASTN	577	1e-39	75
2622	25188	700760643H1	SOYMON015	g619927	BLASTN	701	1e-49	73
2623	25188	701056127H1	SOYMON032	g1899024	BLASTN	649	1e-45	70
2624	27316	701054167H1	SOYMON032	g3087888	BLASTX	177	1e-17	47
2625	27316	701054157H1	SOYMON032	g3087888	BLASTX	177	1e-17	47
2626	488	700682650H2	SOYMON008	g687676	BLASTN	730	1e-52	77
2627	488	700849894H1	SOYMON021	g687676	BLASTN	582	1e-39	76
2628	-GM32703	LIB3051-008- Q1-E1-C12	LIB3051	g1899024	BLASTN	981	1e-76	77
2629	-GM9523	LIB3049-003- Q1-E1-A6	LIB3049	g619928	BLASTX	203	1e-37	64
2630	12693	LIB3051-106- Q1-K1-A9	LIB3051	g619927	BLASTN	459	1e-38	71
2631	488	LIB3040-006- Q1-E1-A12	LIB3040	g687676	BLASTN	622	1e-41	76
2632	488	LIB3053-008- Q1-N1-C6	LIB3053	g687676	BLASTN	597	1e-39	75



2633	488	LIB3055-008-Q1-N1-F5	LIB3055	g687676	BLASTN	559	1e-36	75
2634	488	LIB3053-010-Q1-N1-D8	LIB3053	g687676	BLASTN	514	1e-32	75

# SOYBEAN FRUCTOKINASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
2635	-700834049	700834049H1	SOYMON019	g1915974	BLASTX	112	1e-10	97
2636	-700905716	700905716H1	SOYMON022	g1915973	BLASTN	774	1e-55	77
2637	-700978126	700978126H1	SOYMON009	g1915973	BLASTN	565	1e-38	77
2638	-700983171	700983171H1	SOYMON009	g1915974	BLASTX	96	1e-9	93
2639	-701069652	701069652H1	SOYMON034	g297014	BLASTN	447	1e-27	73
2640	-701118004	701118004H2	SOYMON037	g2102690	BLASTN	440	1e-26	73
2641	-701209270	701209270H1	SOYMON035	g1052972	BLASTN	648	1e-45	79
2642	1174	700832430H1	SOYMON019	g1915973	BLASTN	638	1e-44	81
2643	1174	701101576H1	SOYMON028	g1915973	BLASTN	592	1e-40	79
2644	1174	700754333H1	SOYMON014	g1915973	BLASTN	323	1e-37	80
2645	1174	701004323H1	SOYMON019	g297014	BLASTN	560	1e-37	80
2646	1174	700988192H1	SOYMON009	g1915973	BLASTN	508	1e-33	78
2647	1174	700646337H1	SOYMON013	g1915974	BLASTX	153	1e-30	79
2648	1174	701039647H1	SOYMON029	g1915973	BLASTN	275	1e-12	80
2649	16472	701155250H1	SOYMON031	g1915973	BLASTN	642	1e-50	78
2650	16472	700953304H1	SOYMON022	g1915973	BLASTN	690	1e-48	79
2651	16472	700725996H1	SOYMON009	g1915973	BLASTN	362	1e-28	73
2652	17936	700965277H1	SOYMON022	g2102690	BLASTN	375	1e-42	77
2653	17936	700746240H1	SOYMON013	g2102690	BLASTN	606	1e-41	74
2654	22120	701215393H1	SOYMON035	g2102691	BLASTX	133	1e-11	86
2655	22586	701009695H1	SOYMON019	g2102690	BLASTN	696	1e-49	76
2656	22586	700900731H1	SOYMON027	g2102690	BLASTN	422	1e-26	76
2657	23551	701053585H1	SOYMON032	g2102691	BLASTX	120	1e-9	92
2658	28587	701156878H1	SOYMON031	g2102690	BLASTN	448	1e-33	72
2659	3876	700942858H1	SOYMON024	g297014	BLASTN	705	1e-49	74
2660	3876	701063105H1	SOYMON033	g1052972	BLASTN	679	1e-47	73
2661	3876	700844831H1	SOYMON021	g1915973	BLASTN	466	1e-37	72
2662	5530	700733713H1	SOYMON010	g1915974	BLASTX	156	1e-26	81
2663	5530	701057239H1	SOYMON033	g1915974	BLASTX	176	1e-17	92
2664	5530	700985231H1	SOYMON009	g297014	BLASTN	222	1e-16	79
2665	5805	701010614H1	SOYMON019	g1915973	BLASTN	958	1e-71	80
2666	5805	701003106H1	SOYMON019	g1915973	BLASTN	679	1e-64	81
2667	5805	700748895H1	SOYMON013	g1915973	BLASTN	475	1e-55	83
2668	5805	700892801H1	SOYMON024	g1915973	BLASTN	639	1e-55	80
2669	5805	700891914H1	SOYMON024	g1915973	BLASTN	639	1e-55	81
2670	5805	700962529H1	SOYMON022	g1915973	BLASTN	622	1e-54	82
2671	5805	700869294H1	SOYMON016	g1915973	BLASTN	760	1e-54	80
2672	5805	700986530H1	SOYMON009	g1915973	BLASTN	761	1e-54	80
2673	5805	700661115H1	SOYMON005	g1915973	BLASTN	682	1e-48	78
2674	5805	701041987H1	SOYMON029	g297014	BLASTN	475	1e-45	83
2675	5805	701006803H1	SOYMON019	g1915973	BLASTN	607	1e-41	80
2676	28587	LIB3028-008-Q1-B1-H3	LIB3028	g2102690	BLASTN	900	1e-66	68
2677	5530	LIB3055-004-Q1-N1-H3	LIB3055	g297014	BLASTN	606	1e-39	76

2678	5805	LIB3065-006-Q1-N1-F11	LIB3065	g1915973	BLASTN	954	1e-81	79
------	------	-----------------------	---------	----------	--------	-----	-------	----

# SOYBEAN NDP-KINASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
2679	33331	701108520H1	SOYMON036	g758643	BLASTN	473	1e-31	75
2680	23595	LIB3050-018-Q1-E1-C4	LIB3050	g758643	BLASTN	295	1e-13	76
2681	33331	LIB3040-037-Q1-E1-D6	LIB3040	g758643	BLASTN	413	1e-47	79

# SOYBEAN GLUCOSE-6-PHOSPHATE 1-DEHYDROGENASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
2682	-700869140	700869140H1	SOYMON016	g2829880	BLASTX	164	1e-15	44
2683	-701065174	701065174H1	SOYMON034	g603219	BLASTX	86	1e-9	76
2684	-701130434	701130434H1	SOYMON037	g1197385	BLASTX	189	1e-19	55
2685	-701149522	701149522H1	SOYMON031	g603219	BLASTX	99	1e-8	71
2686	26484	701003905H1	SOYMON019	g1197385	BLASTX	138	1e-15	81
2687	9136	701038169H1	SOYMON029	g603219	BLASTX	139	1e-21	73
2688	9136	700903571H1	SOYMON022	g603219	BLASTX	144	1e-20	81
2689	9136	701045122H1	SOYMON032	g603219	BLASTX	100	1e-13	79

# SOYBEAN PHOSPHOGLUCOMUTASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
2690	-700554424	700554424H1	SOYMON001	g534982	BLASTX	133	1e-25	60
2691	-700556670	700556670H1	SOYMON001	g3294468	BLASTN	355	1e-43	74
2692	-700563871	700563871H1	SOYMON002	g2795876	BLASTX	101	1e-16	75
2693	-700565101	700565101H1	SOYMON002	g3294466	BLASTN	588	1e-40	68
2694	-700566749	700566749H1	SOYMON002	g1814400	BLASTN	475	1e-41	73
2695	-700681382	700681382H2	SOYMON008	g3294467	BLASTX	98	1e-11	48
2696	-700763827	700763827H1	SOYMON018	g3192042	BLASTX	257	1e-29	60
2697	-700865583	700865583H1	SOYMON016	g3192042	BLASTX	134	1e-17	57
2698	-700891379	700891379H1	SOYMON024	g534982	BLASTX	167	1e-15	53
2699	-700942816	700942816H1	SOYMON024	g3294466	BLASTN	636	1e-44	74
2700	-701004954	701004954H1	SOYMON019	g1814400	BLASTN	790	1e-56	78
2701	-701011364	701011364H1	SOYMON019	g534982	BLASTX	284	1e-32	67
2702	-701057063	701057063H2	SOYMON033	g1814401	BLASTX	121	1e-9	60
2703	-701119491	701119491H1	SOYMON037	g1814400	BLASTN	762	1e-54	76
2704	-701149254	701149254H1	SOYMON031	g534982	BLASTX	147	1e-19	52
2705	10032	700988921H1	SOYMON011	g1814400	BLASTN	908	1e-66	80
2706	10032	701136003H1	SOYMON038	g1814400	BLASTN	842	1e-61	78
2707	10032	700953253H1	SOYMON022	g1814400	BLASTN	808	1e-58	77
2708	10032	701103083H1	SOYMON028	g1814400	BLASTN	813	1e-58	78
2709	10131	701104852H1	SOYMON036	g3294466	BLASTN	302	1e-27	74
2710	10131	700970420H1	SOYMON005	g2829893	BLASTX	240	1e-26	56
2711	1180	701125681H1	SOYMON037	g2829893	BLASTX	163	1e-15	82
2712	1180	700559947H1	SOYMON001	g2829893	BLASTX	163	1e-15	82
2713	1180	700556009H1	SOYMON001	g2829893	BLASTX	102	1e-14	87
2714	13262	701006086H2	SOYMON019	g3294466	BLASTN	734	1e-52	75

2715	13262	701137937H1	SOYMON038	g3294466	BLASTN	491	1e-32	71
2716	13262	701004207H1	SOYMON019	g3294466	BLASTN	271	1e-30	75
2717	13262	700904551H1	SOYMON022	g3294466	BLASTN	473	1e-30	75
2718	13262	701014357H1	SOYMON019	g1814401	BLASTX	210	1e-21	83
2719	13262	701146638H1	SOYMON031	g1814401	BLASTX	111	1e-20	80
2720	13262	700833416H1	SOYMON019	g1814400	BLASTN	374	1e-20	74
2721	13262	701148967H1	SOYMON031	g1814401	BLASTX	194	1e-19	82
2722	13262	701156042H1	SOYMON031	g1814401	BLASTX	182	1e-18	67
2723	13262	700943365H1	SOYMON024	g1814401	BLASTX	168	1e-16	76
2724	13262	701105762H1	SOYMON036	g1814401	BLASTX	165	1e-15	83
2725	13262	701038338H1	SOYMON029	g1814400	BLASTN	186	1e-13	77
2726	13262	700645989H1	SOYMON011	g1814401	BLASTX	133	1e-11	78
2727	13262	700868941H1	SOYMON016	g1814400	BLASTN	181	1e-9	80
2728	19312	701121150H1	SOYMON037	g3294468	BLASTN	501	1e-61	79
2729	19312	700742959H1	SOYMON012	g3294468	BLASTN	440	1e-45	83
2730	19312	701135418H1	SOYMON038	g3294468	BLASTN	459	1e-42	79
2731	19312	700979514H2	SOYMON009	g1814400	BLASTN	612	1e-42	78
2732	19883	701133631H2	SOYMON038	g1814400	BLASTN	758	1e-54	75
2733	19883	700970758H1	SOYMON005	g1814400	BLASTN	717	1e-50	77
2734	19883	701153416H1	SOYMON031	g1814400	BLASTN	691	1e-48	76
2735	26278	701214005H1	SOYMON035	g534982	BLASTX	118	1e-8	47
2736	-GM1647	LIB3028-009-Q1-B1-F3	LIB3028	g534982	BLASTX	192	1e-42	57
2737	-GM17162	LIB3055-012-Q1-N1-B3	LIB3055	g1814400	BLASTN	491	1e-29	62
2738	13262	LIB3028-003-Q1-B1-B11	LIB3028	g1814400	BLASTN	1069	1e-80	76
2739	13262	LIB3054-009-Q1-N1-A12	LIB3054	g1814400	BLASTN	612	1e-40	73
2740	13262	LIB3054-009-Q1-N1-A5	LIB3054	g1814401	BLASTX	200	1e-36	73

#### SOYBEAN UDP-GLUCOSE PYROPHOSPHORYLASE

Seq No.	Cluster ID	CloneID	Library	NCBI gi	Method	Score	P-value	% Ident
2741	-700665357	700665357H1	SOYMON005	g1388021	BLASTX	183	1e-18	69
2742	-700674325	700674325H1	SOYMON007	g218000	BLASTN	645	1e-44	72
2743	-700835903	700835903H1	SOYMON019	g1388021	BLASTX	135	1e-11	68
2744	-700841466	700841466H1	SOYMON020	g1388021	BLASTX	115	1e-14	56
2745	-700846570	700846570H1	SOYMON021	g3107930	BLASTN	486	1e-31	70
2746	-700888547	700888547H1	SOYMON024	g3107930	BLASTN	582	1e-39	81
2747	-700973436	700973436H1	SOYMON005	g1212996	BLASTX	132	1e-15	51
2748	-700985779	700985779H1	SOYMON009	g3107930	BLASTN	958	1e-71	83
2749	-700992994	700992994H1	SOYMON011	g1388021	BLASTX	103	1e-10	64
2750	-701061122	701061122H1	SOYMON033	g1388021	BLASTX	129	1e-19	73
2751	-701063465	701063465H1	SOYMON033	g3107930	BLASTN	426	1e-62	82
2752	-701118256	701118256H1	SOYMON037	g3107930	BLASTN	378	1e-31	83
2753	11810	700952705H1	SOYMON022	g3107930	BLASTN	613	1e-54	80
2754	11810	701060568H1	SOYMON033	g3107930	BLASTN	652	1e-45	81
2755	11810	701002783H2	SOYMON019	g3107930	BLASTN	458	1e-43	80
2756	11810	701202674H1	SOYMON035	g218000	BLASTN	331	1e-33	73
2757	11810	700871590H1	SOYMON018	g218000	BLASTN	326	1e-27	76
2758	11810	700747279H1	SOYMON013	g1388021	BLASTX	154	1e-21	75
2759	11810	701014424H1	SOYMON019	g1388021	BLASTX	131	1e-20	84

2760	11810	701039454H1	SOYMON029	g1388021	BLASTX	157	1e-16	80
2761	11810	701054271H1	SOYMON032	g1388021	BLASTX	154	1e-15	71
2762	11810	700955092H1	SOYMON022	g1388021	BLASTX	154	1e-14	71
2763	11810	701107189H1	SOYMON036	g1388021	BLASTX	155	1e-14	72
2764	11810	701107930H1	SOYMON036	g218000	BLASTN	308	1e-14	75
2765	11810	700904384H1	SOYMON022	g1388021	BLASTX	149	1e-13	72
2766	11810	700729516H1	SOYMON009	g1388021	BLASTX	155	1e-13	72
2767	11810	701009325H1	SOYMON019	g1388021	BLASTX	143	1e-12	75
2768	11821	701060627H1	SOYMON033	g218000	BLASTN	253	1e-26	74
2769	11821	701004671H1	SOYMON019	g21599	BLASTX	166	1e-22	77
2770	11821	700964889H1	SOYMON022	g1388021	BLASTX	167	1e-16	67
2771	13178	700562308H1	SOYMON002	g3107930	BLASTN	1198	1e-91	87
2772	13178	701049018H1	SOYMON032	g3107930	BLASTN	1101	1e-82	88
2773	13178	701126215H1	SOYMON037	g3107930	BLASTN	1072	1e-80	88
2774	13178	701211745H1	SOYMON035	g3107930	BLASTN	1038	1e-77	87
2775	13178	700850417H1	SOYMON023	g3107930	BLASTN	1022	1e-76	87
2776	13178	700665292H1	SOYMON005	g3107930	BLASTN	980	1e-72	88
2777	13178	700994009H1	SOYMON011	g3107930	BLASTN	958	1e-71	86
2778	13178	700895203H1	SOYMON024	g3107930	BLASTN	864	1e-68	86
2779	13178	701151725H1	SOYMON031	g3107930	BLASTN	800	1e-66	87
2780	13178	700988803H1	SOYMON011	g3107930	BLASTN	896	1e-65	80
2781	13178	700646581H1	SOYMON014	g3107930	BLASTN	483	1e-62	81
2782	13178	701153726H1	SOYMON031	g3107930	BLASTN	832	1e-60	87
2783	13178	701152333H1	SOYMON031	g3107930	BLASTN	674	1e-56	79
2784	13178	700756960H1	SOYMON015	g3107930	BLASTN	787	1e-56	86
2785	13178	700556901H1	SOYMON001	g218000	BLASTN	772	1e-55	84
2786	13178	701063605H1	SOYMON033	g3107930	BLASTN	566	1e-51	89
2787	13178	701212385H1	SOYMON035	g3107930	BLASTN	390	1e-23	78
2788	13178	700889518H1	SOYMON024	g3107931	BLASTX	131	1e-10	64
2789	17057	700740176H1	SOYMON012	g3107930	BLASTN	798	1e-57	81
2790	17057	700905747H1	SOYMON022	g3107930	BLASTN	511	1e-33	81
2791	1955	701059208H1	SOYMON033	g3107930	BLASTN	1034	1e-77	85
2792	1955	700984109H1	SOYMON009	g3107930	BLASTN	970	1e-72	84
2793	1955	701209482H1	SOYMON035	g3107930	BLASTN	931	1e-68	84
2794	1955	700554847H1	SOYMON001	g3107930	BLASTN	493	1e-66	83
2795	1955	701150363H1	SOYMON031	g3107930	BLASTN	898	1e-66	84
2796	1955	700986014H1	SOYMON009	g3107930	BLASTN	907	1e-66	84
2797	1955	700564270H1	SOYMON002	g3107930	BLASTN	501	1e-65	84
2798	1955	700844253H1	SOYMON021	g3107930	BLASTN	875	1e-64	84
2799	1955	701140892H1	SOYMON038	g3107930	BLASTN	879	1e-64	83
2800	1955	700685893H1	SOYMON008	g3107930	BLASTN	832	1e-60	81
2801	1955	700789732H1	SOYMON011	g3107930	BLASTN	554	1e-57	84
2802	1955	700902418H1	SOYMON027	g3107930	BLASTN	731	1e-52	83
2803	1955	701128306H1	SOYMON037	g3107930	BLASTN	466	1e-46	82
2804	1955	701057973H1	SOYMON033	g3107930	BLASTN	422	1e-43	78
2805	21035	700946288H1	SOYMON024	g3107931	BLASTX	175	1e-17	72
2806	21035	701043539H1	SOYMON029	g3107931	BLASTX	147	1e-13	71
2807	30564	701063642H1	SOYMON033	g3107931	BLASTX	181	1e-25	75
2808	-GM18453	LIB3065-001-Q1-N1-H4	LIB3065	g1212996	BLASTX	68	1e-29	57
2809	-GM32502	LIB3051-013-Q1-E1-A6	LIB3051	g3107931	BLASTX	227	1e-47	51
2810	11810	LIB3030-010-Q1-B1-H12	LIB3030	g21598	BLASTN	1115	1e-84	76

2811	13178	LIB3056-014- Q1-N1-G7	LIB3056	g3107930	BLASTN	1145	1e-99	84
2812	1955	LIB3056-012- Q1-N1-D4	LIB3056	g3107930	BLASTN	856	1e-62	80
2813	30564	LIB3050-003- Q1-E1-D8	LIB3050	g3107930	BLASTN	1078	1e-81	74
2814	30564	LIB3050-010- Q1-E1-D6	LIB3050	g3107930	BLASTN	1050	1e-78	75

669270" 6342660

**\*Table Headings**

**Cluster ID**

A cluster ID is arbitrarily assigned to all of those clones which belong to the same cluster at a given stringency and a particular clone will belong to only one cluster at a given stringency. If a cluster contains only a single clone (a “singleton”), then the cluster ID number will be negative, with an absolute value equal to the clone ID number of its single member. The cluster ID entries in the table refer to the cluster with which the particular clone in each row is associated.

**Clone ID**

The clone ID number refers to the particular clone in the PhytoSeq database. Each clone ID entry in the table refers to the clone whose sequence is used for (1) the sequence comparison whose scores are presented and/or (2) assignment to the particular cluster which is presented. Note that a clone may be included in this table even if its sequence comparison scores fail to meet the minimum standards for similarity. In such a case, the clone is included due solely to its association with a particular cluster for which sequences of one or more other member clones possess the required level of similarity.

**Library**

The library ID refers to the particular cDNA library from which a given clone is obtained. Each cDNA library is associated with the particular tissue(s), line(s) and developmental stage(s) from which it is isolated.

**NCBI gi**

Each sequence in the GenBank public database is arbitrarily assigned a unique NCBI gi (National Center for Biotechnology Information GenBank Identifier) number. In this table, the

NCBI gi number which is associated (in the same row) with a given clone refers to the particular GenBank sequence which is used in the sequence comparison. This entry is omitted when a clone is included solely due to its association with a particular cluster.

### **Method**

The entry in the “Method” column of the table refers to the type of BLAST search that is used for the sequence comparison. “CLUSTER” is entered when the sequence comparison scores for a given clone fail to meet the minimum values required for significant similarity. In such cases, the clone is listed in the table solely as a result of its association with a given cluster for which sequences of one or more other member clones possess the required level of similarity.

### **Score**

Each entry in the “Score” column of the table refers to the BLAST score that is generated by sequence comparison of the designated clone with the designated GenBank sequence using the designated BLAST method. This entry is omitted when a clone is included solely due to its association with a particular cluster. If the program used to determine the hit is HMMSW then the score refers to HMMSW score.

### **P-Value**

The entries in the P-Value column refer to the probability that such matches occur by chance.

### **%Ident**

The entries in the “%Ident” column of the table refer to the percentage of identically matched nucleotides (or residues) that exist along the length of that portion of the sequences which is aligned by the BLAST comparison to generate the statistical scores presented. This entry is omitted when a clone is included solely due to its association with a particular cluster.